HSV-1 ICP4, A MULTIFACETED RNAPOLII TRANSCRIPTION FACTOR

by

Lauren M. Wagner

Bachelor of Arts, The College of Wooster, 2005

Submitted to the Graduate Faculty of

the School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Lauren M. Wagner

It was defended on

July 20th, 2012

and approved by

Michael Cascio, Associate Professor, Bayer School of Natural and Environmental Sciences

Fred Homa, Associate Professor, Department of Microbiology and Molecular Genetics

Paul Kinchington, Professor, Department of Ophthamology

Martin Schmidt, Associate Professor, Department of Microbiology and Molecular Genetics

Dissertation Advisor: Neal DeLuca, Professor, Department of Microbiology and Molecular Genetics

HSV-1 ICP4, A MULTIFACETED RNA POLII TRANSCRIPTION FACTOR

Lauren M. Wagner, PhD

University of Pittsburgh, 2012

ICP4, of Herpes Simplex Virus type 1 (HSV-1) is responsible for activation of viral Early and Late genes, and is necessary for viral replication. ICP4 contains two transactivation domains separated by a DNA binding domain. The complex structure of ICP4 indicates the possible diversity of the cellular and viral proteins it interacts with to function. ICP4 interacts with a variety of transcription complexes to promote RNA Polymerase II mediated transcription. The structural basis for these interactions has not yet been clearly defined. To more closely examine the structural requirements for ICP4 activities, mutants in conserved and degenerate regions of the N-terminus, in the presence and absence of the carboxyl terminus, were examined for effects on viral gene expression. It was found that i) the amino terminal transactivation domain is strictly required for E gene transcription, ii) multiple conserved regions within the N-terminus contribute to transcription, and iii) the amino terminal and carboxyl terminal transactivation domains cooperate to mediate transcription. Affinity purification assays demonstrated that many of the observed defects in transcription probably resulted from the deletion of regions involved in stabilizing TFIID. Complementation analyses demonstrated that TFIID interactions are stabilized by the presence of one functional N-terminal and C-terminal transactivation domain within an ICP4 dimer.

Affinity purification and mass spectrometry were used to determine the complexity of ICP4 mediated interactions throughout infection in addition to the structural requirements provided by ICP4 for these interactions. Mass spectrometry and western blot data indicated

that ICP4 was found in complex with TFIID prior to other components of the transcription machinery including Mediator and TFIIH. Additionally, the amino terminal 774 amino acids were sufficient for interactions with TFIID, Mediator and TFIIH. While ICP4 has previously only been associated with preinitiation complex formation, components of initiation, elongation, mRNA processing, and mRNA export machinery were also found in complexes with ICP4, suggesting that ICP4 functions as a multifaceted RNA PolII transcription factor. Together, the data presented herein provide an understanding of how the structural complexities of ICP4 provide an interface for the formation of transcription complexes. Additionally, a new model for viral transcription is presented.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

HSV: Herpes Simplex Virus

AIDS: Acquired Immunodeficiency Syndrome

ICP4: Infected Cell Polypeptide 4

VZV: Varicella Zoster Virus

HCMV: Human Cytomegalovirus

HHV: Human Herpesvirus

EBV: Epstein Barr Virus

KSHV: Kaposi's Sarcoma Herpes Virus

HIV: Human Immunodeficiency Virus

APP: Amyloid Precursor Protein

Tk: Thymidine Kinase

UL: Unique Long

US: Unique Short

g(B-M): glycoprotein (B-M)

VP: Viral Protein

VHS: Virion Host Shutoff

ICP: Infected Cell Polypeptide

REST: RE1 Silencing Transcription Factor

CoREST: REST Co-repressor protein 1

LSD1: Lysine Specific Demethylase 1

Sp1: Specificity Protein 1

HCF: Host Cell Factor

ts: Temperature Sensitive

Oct1: Octamer binding protein 1

NF1: Nuclear Factor 1

kDa: kiloDalton

NLS: Nuclear Localization Signal

RNA PolII: RNA Polymerase II

IE: Immediate Early

E: Early

L: Late

OriL: Origin L

OriS: Origin S

CCSC: C-Capsid Specific Component

LAT(s): Latency Associate Transcript(s)

ORF: Open Reading Frame

miRNA: MicroRNA

LAP: LAT Promoter

HAT: Histone Acetyl Transferase

HDAC: Histone Deacetylase

HMT: Histone Methyltransferase

HP1: Heterochromatin Protein 1

PRC: Polycomb Repressor Complex

PCR: Polymerase Chain Reaction

qPCR: quantitative PCR

qRT-PCR: quantitative Reverse Transcriptase PCR

Ser2/5: Serine 2/5

CHD: Chromodomain

PHD: Plant HomeoDomain

NFR: Nucleosome Free Region

Brm1: Brahma 1

Brg1: Brahma Relate Gene 1

TCA: Trichloroacetic Acid

MMLV-RT: Murine Lukemia Virus-Reverse Transcriptase

K: Clearing factor or K factor

ssDNA: salmon sperm DNA

min: Minutes

sec: Seconds

EMSA: Electromobility Shift Assay

SDS-PAGE: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Tss: Transcription Start Site

TBP: TATA Binding Protein

TFII (A-H): Transcription Factor II (A-H)

TAF: TBP Associated Factor

Inr: Initiator

BRE: TFIIB Recognition Element

DPE: Downstream Promoter Element

MTE: Motif 10 Element

DCE: Downstream Core Element

PIC: PreInitiation Complex

TFTC: TBP Free TAF Containing Complex

SAGA: Spt-Ada-Gcn5-Acetyltransferase

GTF: General Transcription Factor

CTD: C-terminal Domain

snRNA: Small Nucleolar RNA

CAK: Cyclin Activating Kinase

CBP: Creb Binding Protein

CDK: Cyclin Dependent Kinase

SWI/SNF: SWItch/Sucrose Non Fermantable

DSIF: DRB Sensitivity Inducing Factor

NELF: Negative Elongation Factor

SEC: Super Elongation Complex

CPSF: Cleavage and Polyadenylation Specificty Factor

CstF: Cleavage Stimulatory Factor

PAP: Poly-A Polymemrase

TREX: Transcription Export Complex

ChIP: Chromatin Immunoprecipitation

MTA: 5’-methylthioadenosine

TAP: Tandem Affinity Purification

PKA: Protein Kinase A

vDNA: viral DNA

Arp5: Actin Related Protein 5

SpC: Spectral Counts

RUVBL: RuvB Like

PD: Pull Down

FT: Flow Through

LC-MS-MS: Liquid Chromatography Mass Spectrometry Mass Spectrometry

CBP: Calmodulin Binding Peptide

SBP: Streptaviding Binding Peptide

N-terminal: Amino Terminal

C-terminal: Carboxyl Terminal

PFU: Plaque Forming Units

EAF: ELL Associated Factor

ELL: 11-19 Lysine rich Leukemia Protein

Med: Mediator

Aa: amino acid

Hpi: Hours post infection

Hr: Hour

Wt: Wild Type

PVDF: Polyvinylidine Fluoride

1.0 INTRODUCTION

Infection with Herpes Simplex Virus (HSV) exhibits a variety of pathologies ranging in severity from the common cold sore to life threatening encephalitis. The more severe pathologies associated with HSV infections often occur in immunocompromised individuals. With the increasing number of AIDS patients and the development of resistance to current therapies, studying HSV infection is of the utmost importance. Additionally, HSV is currently being investigated for use as a vector for gene therapy and oncolytic virotherapy for the treatment of a variety of diseases ranging from chronic pain to glioblastoma. The use of HSV in gene therapy and oncolytic virotherapy requires a deep understanding of the virus at multiple levels, including the events regulating the viral transcriptional cascade. Herein, the mechanisms of cellular transcription will be compared to the current understanding of viral transcription, with a particular emphasis on the viral encoded transcription factor Infected Cell Polypeptide 4 (ICP4). The mechanisms utilized by ICP4 to redirect active transcription from the cellular genome to the viral genome will be discussed at length. Additionally, the current gaps in knowledge with respect to viral transcription will be addressed.

1.1 PATHOLOGY AND EPIDEMIOLOGY

1.1.1 Herpesviridae

Herpes Simplex Virus has been documented in literature for at least 2500 years. Hippocrates of Kos initially described the disease as “herpes” Greek for “ to creep or crawl” in reference to the spreading nature of the blisters associated with the disease. The nature of Herpes transmission was not documented until 1500 years later, in 1893, when a French dermatologist, Jean Baptiste Emile Vidal, reported the transmission of HSV between human subjects. Nearly 100 years later, in the 1980s, treatment for HSV infections became available to the public.

Herpes Simplex Virus is a member of the family *Herpesviridae*. Classification into the family *Herpesviridae* is based on the structure of the virion, which consists of a lipid envelope studded with viral glycoproteins surrounding the icosadeltahedral capsid in which the linear double stranded viral DNA is packaged (Figure 1) (231, 242, 336, 337). Between the capsid and the envelope is a proteinaceous layer referred to as the tegument. The tegument contains proteins involved in the activation of viral replication (209). In addition to having similar virion morphology, all members of *Herpesviridae* replicate within the nucleus, destroy the host cell upon the egress of viral progeny, and have the ability to establish a lifelong latent infection within the host (reviewed in ref 278). Currently there are 9 known human herpesviruses with drastically different pathologies, host cell range and replication cycles.

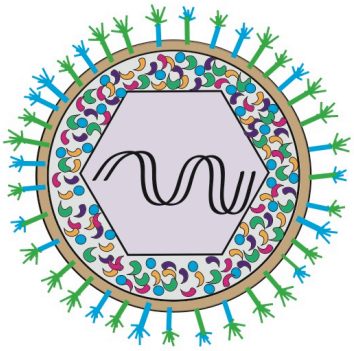


Figure 1: HSV Virion Structure.
Cartoon image of a HSV virion (Figure adapted from 105).

Herpesviridae can be further divided into three general sub-families, *Alphaherpesvirinae, Betaherpesvirinae, and Gammaherpesvirinae*, based on the host cell range, genome configuration, and replication cycle. Members of *Alphaherpesvirinae* include Herpes Simplex Virus (HSV) -1, HSV-2, and Varicella Zoster Virus (VZV). These viruses are characterized by their relatively short replication cycle and wide host range. These viruses establish latent infections within the ganglia of the host. Members of *Betaherpesvirinae* include Human Cytomegalovirus (HCMV), Human Herpesvirus (HHV)-6A, HHV-6B, and HHV-7 and are characterized by a lengthy replication cycle and a limited host cell range. Members of *Gammaherpesvirinae* include Epstein Barr Virus (EBV) and Kaposi's Sarcoma Associated Herpesvirus (KSHV). A limited host range and the ability to establish latency within lymphoid tissues characterize the Gammaherpesviruses.

1.1.2 Herpes Pathology

HSV establishes a lifelong latent infection within the host. Reactivation of latent virus can lead to a symptomatic infection, similar to the primary infection. HSV infection can manifest with a variety of pathologies dependent, partially, on the initial site of inoculation. The most common disease pathology is *herpes labialis,* which manifests as small vesicular blisters presenting on the lips or mouth. Infection of the genitals leads to *herpes genitalis*, with similar vesicular blisters presenting on the genitals. These blisters more readily form ulcers in the genital region due to the thinness of the epithelium in the genital tract. In addition, HSV infection can lead to cutaneous infections such as *herpes whitlow*, an infection of the fingers that is problematic for healthcare workers (280), *herpes gladiatorum,* an infection of the body common amongst wrestlers (153), and *herpes eczema*, a potentially life threatening infection of the skin of eczema patients (35). Ocular infection with HSV can lead to *herpes stromal keratitis*, the leading cause of infectious blindness in the developed world (reviewed in 160). Additionally, infection of immunocompromised individuals, such as infants, AIDs patients, and transplant patients, and less frequently, healthy individuals, can lead to a disseminated infection or encephalitis, which can be fatal if left untreated (279).

In addition to the primary pathophysiology associated with HSV, a relationship between HSV and several other diseases, including atherosclerosis, Alzheimer's Disease, and HIV transmission have recently been suggested (reviewed in 42). Atherosclerosis is a hardening of the arteries and is the leading cause of heart attack. HSV DNA can be found in circulating monocytes and in the coronary arteries of the aorta. With respect to Alzheimer's, HSV DNA has been detected in the brain of Alzheimer's patients and has been shown to affect amyloid precursor protein (APP) processing (42); APP processing leads to the production of β-amyloid,

the major component of amyloid plaques found in the brains of Alzheimer's patients. While a causal relationship between HSV and either atherosclerosis or Alzheimer's has not been established, HSV may contribute to disease progression. HSV infection of the genital tract (generally HSV-2) leads to increased transmission rates (~3 fold) of HIV, due to accessible routes of entry for HIV in the ulcers formed from productive HSV infection (reviewed in 109, 252, 341).

Herpes is a ubiquitous pathogen that affects people globally. Transmission of HSV occurs through physical contact with infectious virus, generally from mucosal membranes. Asymptomatic shedding of the virus frequently occurs, and as such, transmission often goes unnoticed. It is estimated that 31% of children under the age of 13 are infected with HSV-1, with the incidence increasing from about 26% in children 6-7 years of age to 36% in children 12-13 years of age (363). By the age of 50 it is estimated that 70% of the population is seropositive for HSV-1 (289). The CDC reported in 2010 that 16.2% of United States citizens are infected with HSV-2, with the vast majority of infected individuals unaware due to the asymptomatic nature of the virus and a lack of education regarding symptoms. Generally speaking, most disease pathologies associated with HSV can be caused by either HSV-1 or HSV-2, however, HSV-1 is generally associated with oral infections and HSV-2 is generally associated with genital infections. With increasing oral-genital contact arising from evolving sexual practices, the physiological distribution of HSV-1 and -2 is changing (28, 364).

An FDA approved vaccine for the prevention of HSV infection has not yet been developed. The current therapy for HSV infection is treatment with aciclovir, famciclovir, or valaciclovir, which are nucleoside analogues that differ slightly in chemical structure and absorption rates (reviewed in 257). These compounds rely on phosphorylation by thymidine

kinase (51, 89). Once phosphorylated, the viral DNA polymerase incorporates the nucleoside analogue into the replicating DNA causing chain termination (88). Treatment for HSV infection is not generally necessary, given that symptoms will dissipate within 1-2 weeks. However, treatment for HSV infection would help prevent transmission of HSV, thus leading to a reduction in transmission of HIV as well (341). Additionally, severe cases of herpes, such as those associated with ocular infection, *eczema herpeticum*, disseminated infection, and encephalitis do require treatment. Aciclovir resistant strains of HSV have been well documented and do pose a problem, particularly to immunocompromised hosts (102, 257). Other drugs targeting the viral DNA polymerase such as foscaranet and vidarabine do exist, but are not widely used due, in part, to the severe toxicity. The determination of novel drug targets and vaccine strategies to counteract HSV infection is a continuous process.

1.1.3 Herpes as a therapeutic agent

HSV is being investigated for potential medical contributions as a vector for vaccine development, gene therapy, and oncolytic virotherapy. Generally speaking, three different strategies have been employed for the use of HSV as a vector: amplicons, replication defective viruses, and replication competent viruses. Amplicons consist of plasmid DNA carrying the transgene of interest concatamerized and packaged into an empty virus particle. Replication defective viruses have deletions in one or several essential genes, generally involved in transcription and immune surveillance. Replication competent viruses contain deletions within genes that are not essential for replication in vitro, but cause severe attenuation *in vivo*. These include ribonucleotide reductase, thymidine kinase, and the neurovirulence factor γ34.5. Each of these strategies has advantages and disadvantages generally related to transgene expression,

delivery, immune response, and cytopathic effects and have been recently reviewed elsewhere (215).

There are many benefits to using HSV as a vector for vaccine development. Firstly, the dsDNA nature and large genome size make it capable of carrying large heterologous DNA inserts. Additionally, HSV infects a broad host cell range and elicits durable immune responses through the induction of inflammatory cytokines and type I interferons. The use of replication defective vectors allows for efficient delivery of heterologous DNA, elicits an immune response, is noncytopathic, and circumvents the virus's immune evasion strategies (137, 348). Current studies are investigating the potential use of HSV vectors for HIV vaccine development with moderate success.

The neurotropic nature of HSV makes it an ideal candidate for the delivery of transgenes to the nervous system. HSV has been studied for use in a variety of neurological disorders including erectile dysfunction (159), chronic pain(103), diabetic neuropathy(359), and Parkinson's disease(315). The efficacy and duration of transgene expression needs to be further defined in these systems.

In addition to use as a vector for vaccine development and gene therapy, HSV also has emerged as a promising vector for cancer therapy. The use of a virus as an oncolytic vector requires that it be capable of destroying cancer cells while leaving normal cells intact. The very first use of HSV based oncolytic vector was a conditionally replicating mutant deleted for tk (218, 246, 267), which allowed the virus to replicate in cancer cells, but not within neurons or other cells within the brain. Deletions within the neurovirulence factor, γ34.5, in combination with alterations in glycoproteins or deletions of the ribonucleotide reductase, ICP6, have also been investigated as they maintain the intact tk gene, thus maintaining sensitivity to acyclovir.

HSV oncolytic vectors have been used in clinical trials to investigate their effects on glioblastoma, metastatic melanoma, oral squamous cell carcinoma, extracranial solid tumors, colorectal cancer, head and neck cancer, pancreatic cancer, and liver cancer (reviewed in 37, 266). Overall, many of these vectors are promising, showing minimal toxicity to normal cells and frequently extending the lives of patients. Some vectors are currently in stage III clinical trials. The overall efficacy, delivery, and targeting of oncolytic HSV vectors needs to be improved before widespread use as a therapeutic agent can occur.

1.2 HSV LIFECYCLES

HSV transmission occurs through the physical contact of epithelial cells, generally mucosal, with infectious virus. Once the infectious virus has entered the host epithelial cell it undergoes a round of viral replication resulting in the death of the host epithelial cell and the production of progeny virions. The progeny virions are then capable of infecting either nearby epithelial cells or the innervating peripheral neurons. Infection of the peripheral neurons leads to the establishment of a latent infection within the neuron, characterized by the absence of viral protein production resulting from the repressed state of the viral genome that is maintained within the neuron as an episome for the lifetime of the host. Upon certain assaults to the host immune system, the virus reactivates causing symptomatic disease at the site of initial infection.

1.2.1 The viral genome

The viral genome encodes the viral genes necessary to mediate viral infection and replication. Therefore, a brief overview of the viral genome structure will be provided. The dsDNA genome is approximately 152 kb in length and is disproportionately G-C rich (219). The genome encodes approximately 84 gene products and is composed of a unique long (UL) and a unique short (US) region flanked by a series of repeats (Figure 2). The repeated regions on the UL region are termed ab and those on the US region are termed c. The propensity for HSV to undergo homologous recombination and the presence of the repeats within the viral genome lead to four isomers of the viral genome dependent on the orientation of the UL and US regions (127). The presence of repeated regions provides for some genes to be present in two copies within the viral genome (168). Many sequences within the repeated regions are integral to viral replication (168), including coding sequences for Infected Cell Polypeptdes (ICPs) 4, 0, γ34.5, and an origin of replication (OriS).

Figure 2: Structure of the viral genome
The genome contains a UL and US region of the genome that are flanked by inverted repeats.

1.2.2 Lytic infection: Entry through Egress

Infection of host epithelial cells with HSV involves both host cell receptors and viral glycoproteins found on the viral envelope (reviewed in 86). The virus can enter cells through either a direct fusion mechanism or through endocytosis, which is sometimes pH dependent.

Whether the virus enters through a fusion mechanism or through endocytosis appears to be highly cell type specific (64, 225, 241). Viral entry is mediated by multiple viral glycoproteins embedded in the viral envelope, although 4 of them have been described as necessary and sufficient for viral entry: gB, gD, gH and gL (reviewed in 128). gH and gL form a functional heterodimer that will be referred to as gH/gL from here on. Viral entry begins by the binding of gB or gC to heparin sulfate on the cell surface. It is believed that this binding promotes viral entry by tethering the virion to the cell surface. Upon binding of gB to heparan sulfate, gD binds to one of three receptors, nectin-1 (179), herpesvirus entry mediator (HVEM) (230, 354), or 3-O-sulfated heparan sulfate (329). Nectin-1 is the main receptor for cells that are generally infected by HSV including epithelial cells and neurons. Current evidence suggests that nectin-1 is the primary receptor for entry *in vivo*, as in its absence viral infection of neurons is severely attenuated (235). Upon binding of gD to one of its receptors, a conformational change occurs triggering the activation gB and gH/gL proteins and their receptors, leading to membrane fusion (86).

Once the virus has fused with the cell a number of events occur. Firstly, most of the tegument proteins are released into the cellular cytoplasm. Notable tegument proteins include VP16 and virion host shutoff (VHS) (209). VHS is an endoribonuclease that is responsible for the degradation of host cell mRNA, although it can also degrade viral mRNA (186). VP16 is a potent viral transactivator that is carried into the cell in the tegument of the virus (38). VP16, in combination with cellular factors HCF and Oct-1 transcribe the immediate early genes from the viral genome (178). The de-enveloped nucleocapsid containing the viral DNA is transported to the nucleus in a microtubule dependent manner (305), and the viral DNA is injected into the nucleus of the host cell through the nuclear pore (12, 305). It is in the nucleus that viral gene

transcription occurs in an ordered temporal cascade that is initiated by multiple viral and cellular factors.

Viral transcription is a major focus of the work presented in this study, and as such will be discussed at length in a different section of this manuscript. Of importance, RNA Polymerase II transcribes the viral genome (3) in an ordered fashion from the viral genome, which begins by the VP16 activated transcription of Immediate Early (IE) genes, followed by the ICP4 dependent activation of Early (E) genes, viral DNA replication, and ICP4 dependent activation of Late (L) genes (136). Viral mRNA is capped and polyadenylated, however as most viral genes do not contain introns, most viral mRNA is not spliced. The newly synthesized viral mRNA is exported from the nucleus with the aid of the IE protein ICP27 where it is translated by polysomes (reviewed in 286).

Most of the E proteins are involved in DNA replication and nucleotide metabolism. HSV encodes seven proteins essential for viral DNA synthesis. These include UL9 (origin binding protein), ICP8 (ssDNA binding protein), UL30/UL42 (DNA polymerase), and UL5/UL8/UL52 (helicase/primase) (346). The viral genome contains three origins of replications, two within the repeated regions of the genome (OriS) and one within the UL region of the genome (OriL) (311, 312, 352). The exact function of the two different origins is currently not understood, although deletion studies indicate that OriL may play a significant role *in vivo(9)*. During DNA replication UL9 recognizes the origin (either OriS or OriL), and distorts the viral DNA(171). UL9 is necessary for early steps in DNA replication, but appears to be non-essential later in infection (23). Upon UL9 binding, ICP8 also binds and augments the functions of UL9 in addition to distorting and destabilizing the viral DNA (24). UL5/UL8/UL52 next binds, unwinds the DNA and adds a short RNA primer to initiate viral DNA replication (81, 377). The

UL30/UL42 viral DNA polymerase is then capable of initiating viral DNA replication of both the leading and lagging strand (115). Newly synthesized viral DNA is found as in a branched, head to tail concatamer in infected cells (146). Currently, the favored hypothesis is that viral DNA replication occurs via a rolling circle mechanism; however, this has not been proven in infected cells. A recombination dependent mechanism has also been suggested(145), which is similar to that of both T4 and λ phages (reviewed in 346).

Once viral DNA has been synthesized, capsids are assembled and packaged with the viral DNA (reviewed in 59). Capsids are composed of both major and minor capsid proteins. Major capsid proteins include VP5, VP19C, VP23, and VP26 while minor capsid proteins include UL6, UL15, UL17, UL25, UL28, and UL33. The minor capsid proteins are important for processing and packaging the viral DNA into pre-formed capsids and are present to different extents on different types of capsids. Sucrose gradient fractionation from infected cell lysates reveals that there are three types of capsids that differ in density (113). They have been named A, B, and C based on sedimentation densities, with A being the "lightest" and C being the "heaviest". C-capsids are fully packaged, contain viral DNA, and are the type of capsid present in the infectious virion. A-capsids do not contain viral DNA or the scaffolding protein. It is thought that A-capsids result from abortive packaging. Finally, B-capsids contain the scaffolding protein, but do not contain viral DNA (238). It is thought that B-capsids arise from a structural alteration/angularization that prevents them from being packaged. All three forms of capsids arise from an initial procapsid (25, 113, 134), which is the first closed capsid structure that can be visualized in infected cells and contains the major capsid proteins in addition to the scaffold protein.

Once the procapsid is formed, it is packaged with the viral DNA. Packaging of the DNA occurs through a "ring-like portal" that is found at one unique vertex on the capsid (239, 240). There are seven proteins that are necessary for proper packaging of the viral genome. These include the portal protein, UL6, the terminase complex of UL15, UL28, and UL33, UL25, which is important for retention of the viral DNA within the capsid, and UL17 and UL32 whose functions in capsid formation are currently unknown (59). The model for DNA packaging involves the UL end of the genome being inserted into the procapsid at the portal. The terminase complex is necessary for the translocation of the viral DNA into the capsid and for cleavage of the concatameric viral DNA (14, 131). A specific cleavage site has not been identified although it is contained in the a-sequence; it has been hypothesized that cleavage occurs when a "head full" of DNA is obtained, similar to phage (50).

Once a packaged capsid has formed in the nucleoplasm of the cell, it must escape the nucleus and egress from the host cell. Nuclear egress relies on the presence of a UL25/UL17 complex on the capsid and two viral proteins involved in membrane budding, UL31 and UL34 (8). The UL25/UL17 complex is referred to as the C-Vertex Specific Complex (CVSC) (58, 333). It is thought that the CSCV helps ensure that C-capsids are preferentially enveloped. It is hypothesized that the CVSC interacts with two viral proteins, UL31 and UL34, that are found in at the nuclear membrane, and that this interaction helps to direct budding (333). UL31 and UL34 also help to destabilize the nuclear envelope, which is believed to aid in nuclear egress. During infection, the nuclear membrane becomes studded with a variety of proteins that are a part of the mature virion including VP16, VHS, VP22, US3, and a number of glycoproteins (8). When the capsid buds through the nuclear membrane it acquires these proteins, which are associated with

the tegument. Once through the nuclear membrane, the "perinuclear" virion travels through the trans-Golgi network, where it acquires its envelope and is exits the cell via exocytosis.

1.2.3 Latency

Subsequent to lytic infection, the newly synthesized viral progeny escape the host epithelial cells where they obtain access to the surrounding epithelial cells in addition to sensory neurons innervating the site of acute infection. Newly synthesized viral progeny enter neuronal cells through pH-independent fusion events (265). The viral capsid is then transported, in a retrograde fashion, to the nucleus of the neuron (reviewed in 62). For transport to the neuronal cell body, HSV-1 relies on the cellular protein dynein (78). Dynein is a molecular motor that physically moves along the microtubules contained in the axon of the neuron to transport vesicles, organelles, or in this case capsids, to the nucleus of neuronal cells (reviewed in 166). It is hypothesized that multiple viral proteins associated with the inner tegument or the capsid itself, including VP26, UL25, and UL36 are involved in dynein mediated interactions (78, 80, 330). Upon entry into the nucleus of neurons, the viral genome associates with heterochromatin (71, 167, 237, 342) and remains within the neuron as an episome for the lifetime of the host.

During latency the only gene products produced are referred to as the LATs (Latency Associated Transcripts) (310). LAT is produced as a primary 8.3-8.5kb polyadenylated transcript, which is subjected to splicing to yield a stable 2kb intron. The 2kb intron exists in a lariat structure, which stabilizes the transcript from degradation, thus allowing the RNA species to persist for days in infected cells (274, 323, 358). The LAT transcript is expressed abundantly in the nucleus of some, but not all, infected sensory neurons, and is thought to play a large role in the determination of the neuronal cell population in which latency is established (17, 216).

Interestingly, no protein products have been detected from the LAT transcript, despite encoding multiple ORFs. This leads researchers to believe that the effects of LAT are a direct consequence of a function of the DNA and RNA sequences as opposed to a protein product.

Interestingly, while LAT expression is abundant in latently infected nuclei, lytic gene expression is severely inhibited. In fact, it appears as though expression of LAT and lytic genes are mutually exclusive, suggesting a possible role for LAT in the repression of lytic transcription. To this end, LAT has been shown to increase the deposition of repressive heterochromatic marks while decreasing the deposition of transcriptionally permissive euchromatic marks on lytic genes such as ICP4, tk, ICP27, and ICP8 (342). This may provide a potential mechanism for the repression of viral transcription and the establishment of latency.

It has been recently established that HSV-1 encodes for 8 miRNAs that reside within the LAT transcript (61, 334, 335). Interestingly, many of these miRNAs can be found within regions corresponding to ICP0, ICP4, and γ34.5, proteins associated with derepression of the genome, transcriptional activation, and neurovirulence, respectively (277, 334, 335). Mechanistically, these miRNAs may play a role in regulating, or preventing, viral gene expression during latency, by targeting the mRNA of two of the most potent transcriptional activators encoded by the viral genome. While some *in vitro* data supports a model for the regulation of lytic gene expression via these miRNAs, to date there is no evidence that supports a role for these miRNAs in reactivation *in vivo* (234).

While multiple studies have implicated LAT in repression of the viral genome, it does not appear to be necessary to establish latency. Deletions within the LAT promoter (LAP), resulting in ablated expression of LAT, yield viruses that are capable of establishing latency, but are somewhat restricted in induced reactivation (130, 194). These data suggest that expression of

LAT, or a region within the LAP may be involved in activation of the viral genome. However, this phenotype appears to be dependent on both the animal model and type of stress used to induce reactivation (72). Interestingly, recent data has suggested that the 5' exon region of LAT, which was deleted in the aforementioned studies, contains an enhancer element that acts as a binding site for both chromatin repressors and activators. It has been shown that deletion of just this enhancer element, without affecting LAT production, also restricts induced reactivation (22, 210), implicating the enhancer region, not LAT itself in reactivation.

The current model for latency suggests that either activators or repressors bind to the enhancer region of LAT. Whether an activator or repressor binds to this region may be determined by the phenotype of the neuron. Supporting this notion is the fact that only 1/3 of infected neurons actually express detectable levels of LAT, suggesting that there are two transcriptionally different sets of neuronal cells; those that express LAT and those that do not. It has been suggested that in the case of those that do express LAT, an activator binds the enhancer region. To prevent expression of other lytic genes, the promoters of the nearby lytic genes (ICP0) are separated by a highly hetero-chromatinized CTCF insulating region (5). This excessive amount of heterochromatin protects them from activation from the activators bound to the LAT enhancer (5, 180). It is theorized that upon cellular stress, the heterochromatin in the CTCF region breaks down providing the activators bound to the LAT enhancer access to the ICP0 promoter. ICP0 is known as a promiscuous transactivator, based on its ability to alter chromatin to a more derepressed state during reactivation (49, 99-101). As such, ICP0 expression would lead to derepression of the viral genome, thus allowing for transcription of multiple lytic genes, including ICP4, and reactivation to occur.

The current consensus in the field is that ICP0 is necessary to de-repress the viral genome to trigger reactivation. However, it has been shown that while ICP0 is necessary for reactivation, is not necessary for initiating the exit from latency *in vivo* (326). It is instead suggested that VP16, the transactivator of IE gene expression, plays this role. Logistically, VP16 is carried into an epithelial cell within the tegument of the incoming virion to initiate the transcription of IE genes. During latency, VP16, which is expressed with leaky late kinetics, would not initially be present, leading to questions regarding how the coordinated transcription of IE genes occurs. One current hypothesis for the exit from latency suggests that the VP16 promoter is expressed with alternative kinetics during latency, and that it is expressed prior to other lytic genes; thus, VP16 would coordinate the activation of a temporal cascade of transcription during the exit from latency (287, 325). Additionally, it was shown through a variety of loss of function mutations and chemical inhibitors, that ICP0, ICP4, and viral DNA replication are not necessary for expression of VP16 in mouse ganglion and that VP16 was required for the activation of lytic gene expression (325). The question remains though, of how VP16 is induced in latently infected neurons. The authors suggest that there is a neuronal specific repressive factor (perhaps related to the LATs) that must be overcome for VP16 induction, and that VP16 must be produced to sufficient levels to induce IE gene expression.

Interestingly, recent work from the Mohr, Chao, and Wilson labs has shown that reactivation from neuron cultures *ex vivo* is biphasic with an initial global expression of all classes of viral mRNAs, followed by a second phase corresponding to a temporal cascade of gene expression, much like what is seen in lytic infection (164). The authors theorize that there is a decision period following the first phase of gene expression; either the genome can proceed to the second phase of gene expression, leading to reactivation, or the genome can become

repressed once again, presenting a potential molecular switch between latency and reactivation. The first phase of gene expression may correlate to global derepression of the chromatinized viral genome, perhaps as a function of expression of ICP0 (49, 99). Depending on the neuronal and molecular environment, the second phase of gene expression and reactivation would occur. The authors suggest that the localization of the cellular factor HCF may have a pivotal role in this process. Further studies investigating the epigenetics during each phase of gene expression during reactivation would provide clarity into the triggers allowing for progression from phase one to phase two.

In the aforementioned studies, the Triezenberg and Sawtell groups suggest that VP16 is necessary for efficient lytic gene expression in the exit from latency while the Chao, Mohr, and Wilson group suggests that a general derepression of the heterochromatinized genome occurs prior to a lytic cascade of gene expression. Of note, the coordinated second phase of gene expression required the presence and nuclear translocation of a functional VP16 (164). These studies differ partly in the determination of gene expression. The Triezenberg and Sawtell groups determined lytic gene expression based on protein abundance while the Chao, Mohr, and Wilson group use qPCR analyses to determine viral mRNA abundance. Interestingly, in a previous study, the Sawtell group noted the presence of mRNA for many IE genes prior to reactivation stimuli in qPCR analyses, but did not note protein synthesis (326), suggesting, not only a moderately transcriptionally permissive state in their system, but also that viral mRNA processing and translation may play an integral role in reactivation. Perhaps a global derepression of the genome occurs, leading to the transcription of many viral genes. If sufficient viral mRNA export and translation occur yielding appreciable quantities of VP16, IE gene expression would be activated, leading to a further derepression of the viral genome by ICP0 and

the activation of E and L genes by ICP4. It is also likely that the specific quantities of viral proteins produced as well as the availability and localization of certain cellular proteins, such as HCF, play an important role in determining whether a reactivation event occurs.

Once the full exit from latency occurs, viral progeny are assembled and released. The production of viral progeny within the neuron is slightly different than in a normal epithelial cell (see section on lytic phase). In an epithelial cell, the capsid is packaged within the nucleus, travels through the nuclear membranes, obtains the tegument, and becomes enveloped as it traverses through the trans-Golgi apparatus (reviewed in 222). The assembled virion is then transported along microtubules toward the periphery of the cell where they are released via exocytosis. In contrast, in neuronal cells, the virus particles must travel in an anterograde fashion, down the lengthy axon to the tips, or growth cones, of the neuron. Similar to retrograde transport, the newly released viral particles utilize a molecular motor, this time kinesin, to travel the distance of the axon. It is believed that the mature virion does not fully assemble until it reaches the growth cones of the neuron. In contrast to assembly in epithelial cells, the capsid acquires the tegument layers and travels separately from glycoprotein components of the envelope to the growth cones. It isn't until both components reach the growth cones that the capsid acquires the envelope and is released, likely via exocytosis (132, 189-191).

1.3 CELLULAR TRANSCRIPTION

HSV-1 transcription is a tightly regulated, efficient process that, because it utilizes many cellular proteins, has many aspects that are similar to cellular transcription. However, the virus alters the cellular transcription machinery such that it has a higher affinity for viral activities compared to

cellular ones. As such, a thorough understanding of cellular transcription is necessary to examine how viral infection exploits cellular mechanisms to transcribe the viral genome, leading ultimately to efficient replication of the virus concomitant with cellular death.

1.3.1 Chromatin control of cellular gene expression

Transcription is a highly dynamic, tightly regulated process. The sheer size of the human genome requires it to be packaged and compressed into chromatin. The core unit of chromatin is the nucleosome, which is made up of an octamer of histones: two H2A/H2B dimers and an H3/H4 tetramer (173, 355). Each nucleosome contains 147 base pairs of naked DNA wrapped around each histone octamer 1.65 times, thus leading to compaction of the DNA (213). Many genes are only expressed upon intrinsic or extrinsic stimulation, and as such, the host cell must be capable of maintaining the human genome in a state such that it is compressed into chromatin but retain the ability to quickly and efficiently activate or repress some genes in response to certain stimuli. In this respect, epigenetic modifications of the host genome, chromatin remodeling, and nucleosome spacing play a large role in the repression and activation of gene expression, and are generally thought of as one of the first levels of regulation related to gene expression.

Activation and repression of transcription are intimately linked to the state of epigenetic modifications of the histones. The histones within a nucleosome are structured such that a terminus of each histone is protruding from the nucleosome; this is referred to as the histone "tail". Each histone tail can undergo a variety of reversible post translational modifications including lysine acetylation, lysine and arginine methylation, serine phosphorylation, lysine ubiquitination, and multiple others (196). These covalent modifications are carried out by a

variety of cellular enzymes; histones are methylated by methyltransferases (HMTs) and methyl groups are removed by demethylases, they are acetylated by histone acetyltransferases (HATs) and deacetylated via histone deacetylases (HDACs), and so on. These covalent modifications are then converted into a "histone code" with certain types of modifications and combinations therein, leading to activation or repression of the surrounding DNA (147). The "histone code" is incredibly complex and not fully understood to date.

The covalent modifications of histone tails, in addition to other factors, ultimately leads to either transcriptional permissiveness or repression. Based on the histone tail modifications and the transcriptional state, chromatin can be classified into two general groups, euchromatin and heterochromatin (For a review see 39, 116). The more compact form of chromatin, referred to as heterochromatin, is electron-dense when visualized under EM, and can be further divided into facultative and constitutive heterochromatin based on the reversibility of the histone tail modifications. Constitutive heterochromatin is stable and is generally present at the centromere and telomeres of cellular chromosomes. Constitutive heterochromatin is marked by tri-methylation of H3K9, which is involved in recruiting repressive complexes such as heterochromatin protein 1 (HP1) (196). HP1 binds to methylated H3K9 via its chromodomain, and recruits HMTs to methylate the adjacent histones. It is in this manner that the constitutive heterochromatic mark can "spread" along the DNA. Additionally, it is believed that HP1 acts as a scaffold for the recruitment and binding of other proteins that are involved in maintaining and spreading of heterochromatin. Chromatin boundary elements encoded in the cellular DNA often prevent the spread of these heterochromatic marks (314). While constitutive heterochromatin was originally believed to be irreversible, it has been shown that during cell division, constitutive

heterochromatic marks are lost as a result of a family of demethylases containing a JmjC domain (193, 353).

Differing from constitutive heterochromatin, facultative heterochromatin can be readily reversed into euchromatin and is generally associated with tissue specific and developmental gene expression (For a review of facultative heterochromatin see 332). Facultative heterochromatin is marked by hypoacetylation of histones and the tri-methylation of H3K27 by the polycomb repressive complex (PRC) 2. Trimethylation of H3K27 is recognized by PRC1, which binds to the methylated histones via a chromodomain, and catalyzes the mono-ubiquitination of H2AK119, halting RNA PolII mediated transcription (288). Ubiquitination of H2AK119 also increases its affinity for the linker histone, H1, which results in increased compaction. It is currently believed that non-coding RNAs may be involved in targeting the PRC to DNA to establish facultative heterochromatin. Interestingly, H3K4me3 marks, which are generally associated with active euchromatin, have been found at sites associated with facultative heterochromatin; a "poised" RNA PolII has been found at these sites also. The poised RNA PolII is phosphorylated on Ser5, but not Ser2, indicating that it is capable of initiating transcription, but not elongating. It is postulated that the monoubiquitination of H2A, along with the physical presence of PRC1 prevents RNA PolII elongation, and that deubiquitination and demethylation, along with dissociation of the PRC and the addition of activating marks such as histone acetylation lead to activated transcription within these "bivalent" genes (288).

The transcriptionally permissive state of chromatin is referred to a euchromatin and is characterized by the association of acetyl residues on H3K9 and H3K14 and methyl residues on H3K4 (39). The addition of the acetyl group onto the positively charged lysine group results in the negation of the positive charge, which leads to weakened electrostatic interactions between

the lysine and the negatively charged DNA, and causes a more relaxed, or open, chromatin conformation. Additionally, histone modifications, such as acetylation, can lead to recruitment of chromatin remodeling complexes, which as discussed below, alter nucleosome positioning to make DNA more accessible to DNA binding proteins such as transcriptional activators. Together, the modification patterns of histones provide a complex pattern of regulation by providing binding surfaces for a variety of different transcriptional complexes.

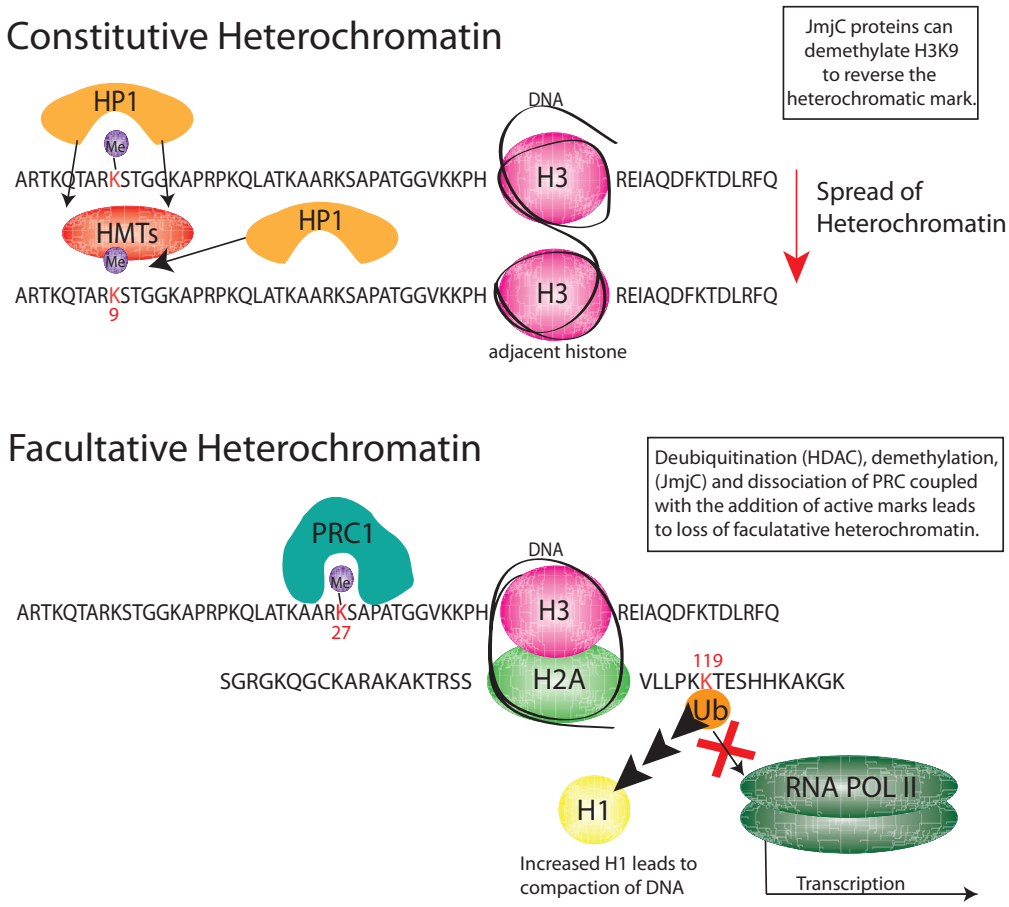


Figure 3. Chromatin States.
Nucleosomes consist of DNA wrapped around a histone core that is composed of two H2A/H2B dimers and an H3/H4 tetramer. Histone tails can be modified via acetylation (Ac), methylation (Me), or ubiquitination (Ub). The histone modification pattern is different for constitutive heterochromatin, facultative heterochromatin, and euchromatin. A simplistic version of the different types of chromatin are displayed in this figure.

Histone tail modifications, in combination with cellular transcription factors, help to direct chromatin remodeling complexes to activated genes (47, 321). There are four general

classes of chromatin remodeling complexes that use ATP hydrolysis to slide, move, or evict histones from the DNA making it accessible to transcriptional regulatory proteins. The four classes of ATP dependent chromatin remodeling complexes include SWI/SNF, ISWI, Chromodomain (CHD), and Ino80, which differ in the architecture of the ATPase and in the composition of the non-catalytic subunits, leading to differing biological functions. Chromatin remodeling complexes generally contain bromodomains, chromodomains, or PHD finger domains allowing for targeting to acetylated or methylated histones. As discussed below, many promoters are free of nucleosomes, and as such chromatin-remodeling complexes likely play a role in exposing distal activator sites, increasing the size of nucleosome free regions, and ensuring proper elongation through chromatinized regions (47).

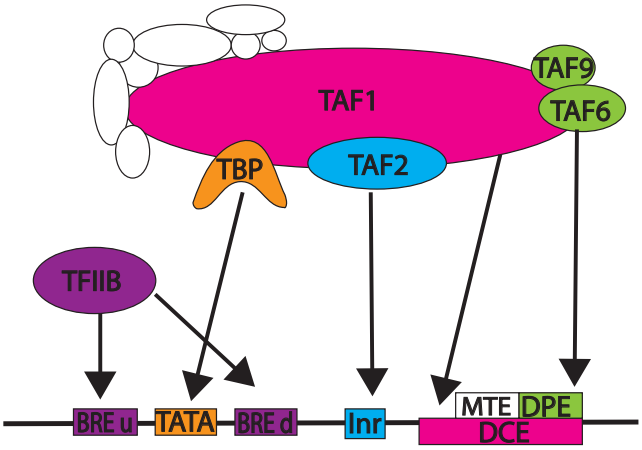
The physical localization of histones on the DNA template plays an important role in the determination of the transcription state of the genome. With the advent of ChIP on ChIP and ChIP-seq technologies, the examination of histone modifications and their deposition along the entire genome has been possible. Evidence from these studies revealed that there is a nucleosome free region (NFR) located near the transcription start site of most active genes. Additionally, there are specifically placed histones both upstream (-1) and downstream (+1) of the NFR, which may partially dictate the transcription start site (10, 369). Generally, the nucleosomes surrounding the transcription start site contain covalent modifications or histone variants that are associated with activated transcription, and these modifications and variants decrease further from the transcription start site (149, 150, 199). It has been shown that general transcription factors and RNA PolII are often bound within the NFR and likely act as the nucleation site for the transcription machinery. Recently it has also been demonstrated that RNA PolII pausing near the transcription start site, prior to initiation, may play a significant role in

regulating the activation of transcription. This pausing may direct ATP dependent chromatin-remodeling complexes to evict or slide the +1 nucleosome allowing for proper initiation to occur (150, 281). Together, histone post-translational modifications, the association of chromatin remodeling complexes, and nucleosome positioning provide a platform for the regulation of transcription within the host cell.

1.3.2 Promoter elements

Transcription initiation begins with cellular transcription machinery binding to promoter regions. In general, promoters can be categorized into two groups, focused and dispersed (reviewed in 156). Focused promoters contain well-defined core promoter elements for transcription factor binding at defined distances from the transcription start site. Dispersed promoters contain multiple transcription start sites within approximately 50-100 bp of DNA, tend to lack many of the core promoter elements found in focused promoters, and are generally found within CpG islands. Interestingly, focused promoters have been well defined in a variety of systems, however, only approximately 1/3 of mammalian promoters are classified as focused. Focused promoters are more highly conserved amongst the evolutionary chain; all HSV promoters studied to date are focused. However, the vast majority of mammalian promoters can be defined as dispersed and as such, have no definable promoter elements or transcription start site. It has been postulated that nucleosome positioning, epigenetic marks, chromatin remodeling factors, and general transcription factors such as TFIIB and TFIIF play a substantial role in defining the transcription start site from dispersed promoters, whereas core promoter elements provide this function on focused promoters.

There are a variety of promoter elements (also called cis-regulatory elements) varying in sequence that promote the binding of specific transcription factors to focused promoters. While only ~10% of cellular promoters contain a TATA box, it remains the most widely studied promoter element. The TATA box is located approximately 35 bp upstream of the transcription start site and has a general consensus sequence of TATAAA. The transcription factor TBP binds to the TATA box and directs general transcription factors, including TBP associated factors (TAFs), and RNA PolII to the promoter region. Another common promoter sequence is the Initiator (Inr) element which has a consensus sequence of YR in mammals and YYANWYY in humans, where Y=pyrimidine, R=purine, N=any nucleotide, and W=A or T. The Inr is generally located at the transcription start site. TAF1 and TAF2 of TFIID have been shown to bind cooperatively to the Inr to activate transcription (43). Additional elements include the BRE (TFIIB Recognition Element), which as its name implies, is bound by TFIIB (187), the DPE (Downstream Promoter Element), which is recognized by TAF6 and TAF9 (33), the MTE (Motif 10 Element), which are also recognized by TAF6 and TAF9 (322), and the DCE (Downstream Core Element), which is recognized by TAF1 (192). Figure 4 displays many core promoter elements, their consensus sequences, localization relative the transcription start site, and cognate-binding partners(324). Many of the core promoter elements function together to activate transcription. For example, both the DPE and the MTE function with Inr sequences to enhance the affinity of TFIID to promoter DNA (34, 203). The exact positioning, function, and cooperativity between core promoter elements and chromatin modifications, and their effects on the activation of gene expression is still largely undefined and consequently, widely studied. Nonetheless, it is evident from multiple genetic experiments on focused promoters, that core promoter elements are vital to directing activated gene expression.



Core Promoter Element	Consensus Sequence (5' to 3')	Bound Protein
BRE u	(G/C)(G/C)(G/A)CGCC	TFIIB
TATA	TATA(A/T)A(A/T)(A/G)	TBP
BRE d	(G/A)T(T/G/A)(T/G)(G/T)(T/G)(T/G)	TFIIB
Inr	PYPYAN(T/A)PYPY	TAF1/TAF2
MTE	C(G/C)A(A/G)C(G/C)(G/C)(G/C)AACG(G/C)	na
DPE	(A/G)G(A/T)CGTG	TAF6/9
DCE	CTTC; CTGT; AGC	TAF1

Figure 4: Promoter elements, consensus sequences, and binding partners.
This figure was adapted from (324).

1.3.3 Pre-Initiation Complex Assembly

Many of the core promoter elements in a focused promoter serve to promote the establishment of RNA PolII and associated general transcription factors on promoter sequences, which is referred to as pre-initiation complex (PIC) assembly. The PIC is composed of TFIID, TFIIA, TFIIB, TFIIF, RNA PolII, TFIIE, and TFIIH. The favored model for PIC assembly

involves a stepwise assembly of general transcription factors and coactivators to a promoter. This stepwise model was established *in vitro* by performing native gel electrophoresis and DNAse footprinting assays using sequential additions of partially purified general transcription factors and RNA PolII (32). These assays and others that followed, suggest that TFIID binds to the promoter first and is stabilized by the addition of TFIIA, then TFIIB. This drives RNA PolII and TFIIF to bind to the promoter, followed by TFIIE and TFIIH.

PIC formation on promoters has been described as the initial step in transcriptional activation. Binding of TFIID to promoter regions nucleates PIC formation. TFIID is composed of TATA Binding Protein (TBP) and 12-14 TBP Associated Factors (TAFs) (104). TBP binds specifically to TATA boxes in the minor groove of DNA, thus providing interaction with the promoter DNA. TAFs have many functions including stabilizing the structure of TFIID, interacting with other core promoter elements, and acetylating histones. Activators, such as VP16 and Gal4 (and many others) interact with different surfaces of the TAF components of TFIID to direct it to the promoter regions. Within the past 20 years it has been noted that complexes that contain many TAFs, but lack TBP exist within the nuclear environment. These complexes are referred to as TBP Free TAF Containing complexes (TFTC) and the best characterized is the SAGA complex (2).

The SAGA (Spt-Ada-Gcn5-Acetyltransferase) complex shares many features with TFIID, although it has different promoter specificity; some genes are TFIID dependent, while others are SAGA dependent (For review see 2, 18, 308). SAGA is composed of multiple subunits including the kinase Tra1, the HAT Gcn5, and TAFs 5,6, 9, 10, and 12. Of note, these TAFs are also considered part of the "core complex" of TFIID (356). Just like TFIID, SAGA contains features that allow for interaction with a variety of transcriptional activators, bind to

histones, and modify chromatin. Interestingly, while SAGA does not contain TBP, different subunits of the SAGA complex are required for the recruitment of TBP at SAGA dependent promoters (229, 299). SAGA can also enhance RNA PolII recruitment by interacting with GTFs such as TFIIA (347). Interestingly, the Mediator and certain GTFs are required for recruiting the SAGA complex to SAGA dependent promoters. Conversely, many GTFs and Mediator are dispensable for recruiting TFIID to promoter regions, but are required for activated transcription of TFIID dependent genes. Thus, there are apparent differences in the recruitment of certain GTFs in PIC formation on SAGA dependent versus TFIID dependent promoters. While SAGA mediated activities are integral for proper cellular transcription, ICP4 mediated transcription from the HSV-1 genome occurs in a TFIID dependent manner, and as such, will be the focus of this discussion.

After TFIID has been recruited to the promoter DNA, TFIIA and TFIIB act to stabilize TFIID on the promoter. Immunodepletion assays indicate that TFIIA is not necessary for PIC formation at all promoters, although the ability of TFIIA to interact with TBP, various other TAFs, activators, and coactivators suggest that TFIIA may be able to enhance TFIID binding at some promoters, and may increase the rate at which PIC formation occurs (157, 201, 247). In the presence or absence of TFIIA, TFIIB binds to the promoter region next. It is thought that TBP, TFIIB, and DNA form a stable tripartite complex that helps stabilize TFIID on the promoter DNA (243). Additionally, TFIIB has been shown to form contacts with TFIIF(122) and RNA PolII (44, 122).

Once TFIIB has entered the PIC, TFIIF and RNA PolII enter the complex based on affinities for TFIIB and each other (264, 306). TFIIF is responsible not only for recruiting RNA PolII to promoters (106), but also for causing a structural change in the promoter DNA, thereby

increasing RNA PolII's affinity for the promoter region (272). TFIIF has many functions related to transcription initiation and elongation, including in transcription start site selection (112), promoter escape (367), and facilitating in RNA PolII elongation with TFIIS to decrease promoter proximal pausing (264, 374). Additionally, TFIIF binding to the PIC is necessary for recruitment of TFIIE and TFIIH to form the fully functional PIC.

TFIIF recruitment coincides with RNA PolII recruitment to promoter regions. RNA PolII is the functional cellular DNA dependent RNA polymerase that transcribes mRNA. RNA PolII consists of 12 highly conserved subunits, RPB1-12 (6). RPBs 1 and 2 contain the majority of the enzymatic activity of RNA PolII as they are necessary for phosphodiester bond formation. RPB1 contains a motif referred to as the C-terminal Domain (CTD), which consists of a set number of heptapeptide (YSPTSPS) repeats. The number of heptapeptide repeats is dependent on the complexity of the organism; humans have 52 (324). The CTD of RNA PolII contains three serine residues, lending it susceptible to phosphorylation by a number of kinases. Generally speaking, serine 2 and serine 5 are heavily phosphorylated and are associated with recruiting 3'modifying enzymes and promoter clearance, respectively (172). More recently, a role for phosphorylation of serine 7 in transcription of snRNA has been established (165). A kinetic relationship between CTD phosphorylation and transcription elongation has also been established, in that serine 5 is heavily phosphorylated near the 5' end of the gene whereas serine 2 phosphorylation is more prominent near the 3' end of the gene (31). Presumably, the complex phosphorylation patterns amongst three serine residues present 52 times on the CTD of RNA PolII leads to varying patterns of gene regulation that are in the process of being defined (84).

Once RNA PolII and TFIIF have bound to promoter sequences, TFIIE is recruited. TFIIE recruitment to the promoter is stabilized by interactions with various other components of

the PIC, including TFIIB, TFIIF, and RNA PolII. Additionally, TFIIE is required for TFIIH recruitment to the PIC (324). TFIIE and TFIIH are both necessary for promoter melting, as evidence by the fact that they are both dispensable for transcription initiation from premelted templates (133, 248). TFIIE is also capable of enhancing many of TFIIH's enzymatic activities.

TFIIH, like many of the GTFs, is composed of multiple subunits conferring many enzymatic functions including ATPase, helicase, and kinase activities (55, 57). The ATPase and helicase functions of TFIIH are important for promoter melting (forming the "transcription bubble"), permitting catalysis of the first phosphodiester bond in the nascent RNA, and for promoter clearance. The helicase functions are involved promoter clearance while the ATPase functions are involved in formation of the transcription bubble (85). The kinase domain of TFIIH, also referred to the CAK, contains cyclin H and the cyclin dependent kinase 7 (CDK7), which is involved in phosphorylating serine 5 of the CTD of RNA PolII (211). The Mediator complex enhances the catalytic activities of the CDK7 homologue in yeast. Ser5 phosphorylation is important for recruitment of the 5' capping complexes and for promoter clearance (172). Many studies have shown that a hypophosphorylated form of RNA PolII remains at the promoter region while a hyperphosphorylated form of RNA PolII can be found downstream of the promoter, suggesting that phosphorylation helps to drive promoter clearance.

1.3.4 Activators and Coactivators

Activated transcription relies not only on PIC formation on promoters, but also on activators and co-activators (40, 276). Activators bind to specific DNA sequences, generally upstream, of the transcription start site, and make contacts with either co-activators or directly with components of the PIC. The general function of activators is to increase transcription rates.

This is performed via a variety of functions including stabilizing PICs, increasing the affinity of the PIC to promoter DNA, or enhancing the kinetics of re-initiation and promoter clearance. Additionally, under certain conditions, activators can also act as repressors, thus increasing the complexity of gene regulation. There are a wide variety of transcriptional activators that are activated in response to a plethora of external or internal stimuli, and as such offer a means to quickly activate, or repress gene expression under specific conditions (40). The sheer number of activators and the varying mechanisms used to enhance, or repress, transcription under various physiological conditions prevents a comprehensive overview of activator function within this manuscript.

Activators often function in the presence of a co-activator. Co-activators generally act as a bridge between the PIC and the activator. Just as with activators, wide arrays of co-activators are utilized to activate transcription (123). Some of the more widely studied co-activators include p300, CBP, SWI/SNF, SAGA, and Mediator. Co-activators usually contain a diverse array of proteins with varying enzymatic functions; this provides for many co-activators to be involved in multiple steps beyond transcription initiation. While the importance of these co-activators has been demonstrated in a variety of systems under varying conditions, based on the focus of the research presented herein, only the Mediator complex will be discussed.

The Mediator complex is a multi-protein complex comprised of four general regions or modules, the head, middle, tail, and a dissociable kinase domain. There are approximately 30 proteins that make up the Mediator complex, which allows for interactions with a vast number of activators, some of which are highlighted in Figure 5 (174, 214). The majority of activators function through the tail module, although interactions with the head and middle region are not uncommon. The ability of Mediator to interact with a wide array of activators is suggestive of

the broad range of genes whose activation are mediator dependent. The importance of the mediator complex in development and transcription is highlighted by it's necessity for yeast viability (236). Additionally, the protein complexity of Mediator provides for functionality at a variety of levels in transcription including chromatin remodeling, PIC formation, promoter clearance, and the transition between productive initiation and elongation; many of these functions are interdependent on similar factors, varying activators, and each other (214).

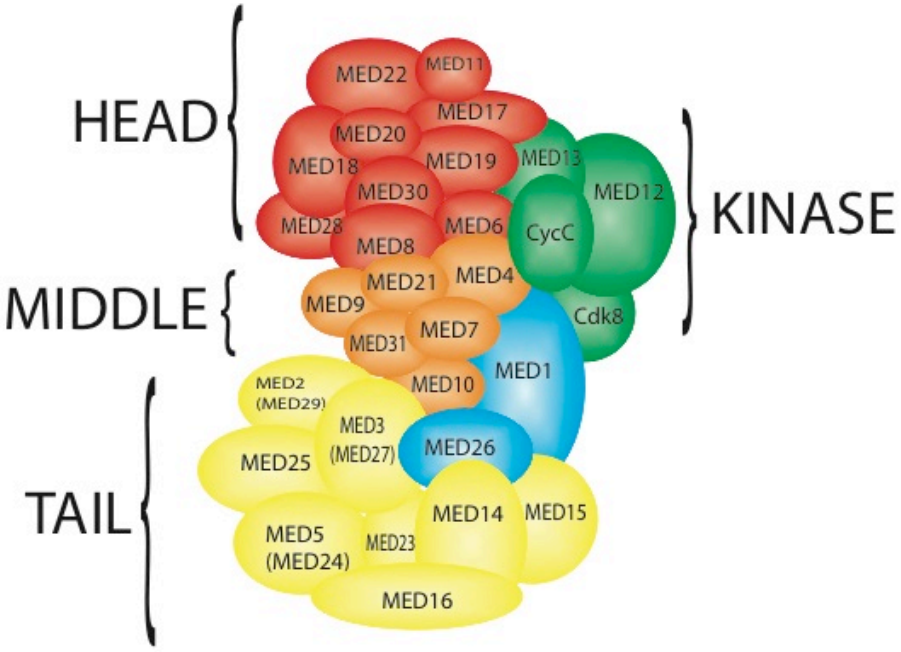


Figure 5: Mediator Complex.
A representative structure of the mediator complex displays the four modules that compose the complex. This figure was adapted from (214).

1.3.5 Elongation and Termination

Once functional pre-initiation complexes have assembled on activated promoters and the initial phosphodiester bonds have been formed, RNA Pol II disassociates from the majority of the initiation machinery and travels along the DNA, elongating the nascent transcript. A number

of molecular events occur that drive the dissociation of RNA PolII and productive elongation. The kinase domain of TFIIH phosphorylates the serine 5 residue of the CTD of RNA PolII, which leads to the recruitment of elongation and processing factors, such as the 5' capping enzyme (172). The additional phosphorylation of TFIIB by TFIIH, also likely plays a role in promoter escape (343). DSIF (DRB Sensitivity Inducing Factor) and NELF (Negative Elongation Factor) play a significant role in promoter proximal pausing. It has been demonstrated that in drosophila, DSIF interacts directly with nascent RNA to prevent elongation (227). Phosphorylation of DSIF, NELF, and Ser2 of the RNA PolII CTD by CDK9 of pTEFb (positive transcription elongation factor B), causes the dissociation and alleviation of the repressive effects (206, 338, 365). Once promoter clearance has occurred, elongation factors such as TFIIS, elongin, the super elongation complex (SEC), and TFIIF increase the processivity of RNA PolII by preventing its dissociation from both the DNA template and nascent RNA (268). Additional chromatin remodeling factors also travel with the elongating RNA PolII, functioning to evict or move nucleosomes in the path of the elongating complex (15). It should be noted that during elongation, splicing complexes are associated with the nascent RNA and promote the specific removal of introns from the transcribing RNA. As most HSV genes are intronless, the mechanisms of spliceosome assembly, RNA recognition, cleavage, and ligation will not be reviewed. For a recent review please refer to (154) and references therein.

Once the elongation complex reaches the poly-A signal (AAUAAA), the newly synthesized transcript is cleaved and exported from the nucleus. Just as RNA PolII phosphorylation of serine 5 leads to the recruitment of capping enzymes, phosphorylation of the serine 2 residue leads to the recruitment of the cleavage and polyadenylation specificity factors (CPSF), the cleavage stimulatory factor (CstF), and the associated poly-A polymerase (1). The

recruitment of these complexes to elongating RNA PolII causes RNA PolII pausing and an endonucleolytic cleavage of the nascent RNA (79). The capped, spliced, and polyadenylated mRNA is then transported through the nuclear pore complex to the ribosomes where translation occurs. At the basic level, export through the nuclear pore is dependent on the TREX complexes, the RNA adaptor protein Aly/Ref, and the mRNA export receptor Tap-p15, in addition to the nuclear pore complex itself (158). Nuclear export and transcription are coupled through a variety factors that interact with both the nuclear export machinery and chromatin remodeling, initiation, or elongation factors.

As mentioned previously, it is believed that much of the initiation machinery, including TFIID, TFIIB, TFIIA, TFIIH, and Mediator remain at the activated promoter, thus allowing for efficient reinitiation of transcription (370). It has been demonstrated that the CPSFs and CstF can interact with TFIIB causing the promoter and terminator regions of the gene to be in close proximity, a phenomena referred to as gene looping. Gene looping functions to enhance the reinitiation kinetics of RNA PolII, thus increasing overall transcription rates (301). Additionally, gene loops have been found at or near nuclear pore complexes, thus also coupling efficient mRNA export to transcription (320). Figure 6 is a representative model of gene looping and depicts the reinitiation scaffold interacting with components of the termination machinery.

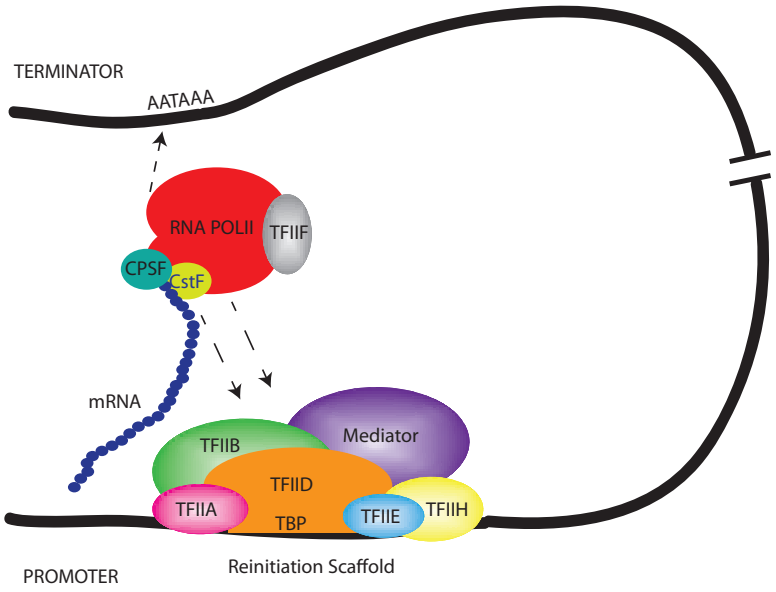


Figure 6: Gene Looping.
During activated transcription a reinitiation scaffold remains at the promoter after RNA PolII has "escaped" the promoter. The reinitiation scaffold can act as a "bookmark" for recycling of RNA PolII. Additionally, components of the reinitiation scaffold interact with components of the termination machinery such as the CPSF and CtsF. This forms the "gene loop" which puts the promoter region in close physical proximity to the terminator region of the gene, enhancing reinitiation kinetics. Figure adapted from (292).

1.4 VIRAL TRANSCRIPTION

Viral gene expression has many of the same characteristics and mechanisms of cellular transcription including chromatin remodeling, PIC formation, initiation, elongation and export. However, there are many notable differences, which will be discussed. Additionally, the viral genome is transcribed in a temporally ordered fashion. The molecular mechanisms underlying viral transcription and its kinetics will be reviewed.

1.4.1 Chromatin states of the viral genome

Prior to the activation of cellular gene expression, the cellular chromatin must be remodeled such that the DNA is accessible to the cellular transcriptional activators and RNA PolII. Similarly, activation of viral transcription also requires remodeling of the chromatinized viral genome to a more permissive state. The chromatin state of the viral genome is largely responsible for the ability of the virus to remain latent; the condensation of the majority of the viral genome into heterochromatin prevents the transcription of viral lytic genes. Reactivation requires a drastic remodeling of chromatin to allow a more permissive state and viral transcription to ensue. Chromatin remodeling also appears to have a role in lytic infection of epithelial cells as well. Because the chromatin state of the genome directly influences the activation of transcriptional processes, the differences between the chromatin states of the viral genome during lytic and latent infection will be discussed.

During lytic infection, viral DNA, much like cellular DNA is associated with histones in infected cells. Evidence has shown that histones are absent from the encapsidated viral DNA within the virion, indicating that histones are derived directly from the host cell upon entry into the nucleus (209, 245). However, unlike cellular DNA, viral DNA is not associated in a normal nucleosomal structure during lytic infection. Initial studies using micrococcal nuclease digestion on DNA isolated from HSV-1 infected cells demonstrated that the viral DNA did not form a laddered pattern, as cellular DNA did, thus suggesting that a regular distribution of nucleosomes did not exist (161, 195, 232). ChIP studies have shown that incoming viral DNA associates with histone H3 quickly (within an hour), but that levels of H3 association decline between 3-6hpi, coincident with, but independent of viral DNA replication (49, 245). Additionally, H3 can be found at each temporal class of promoters (49, 245) and lytic promoters are associated with

euchromatic marks, such as methylation of H3K4, and acetylation of H3K9 and H3K14 (129, 139, 161). The relevance of histone deposition on the viral genome during lytic infection is not fully understood, however, treatment of cells with MTA, a general methyltransferase inhibitor, resulted in a decrease in viral transcription and replication, indicating histone methylation may play an important role in the viral lifecycle (139).

The initial deposition of histones on the viral genome may be a cellular defense against infection. However, as evidenced by the reduction in H3 present on viral genomes later in infection, the virus has likely developed methods to counteract this cellular defense (49, 245). VP16 expression has been shown to reduce the amount of histone H3 deposition on IE promoters to promote IE gene expression. It was demonstrated that the acidic domain of VP16 was responsible for the recruitment of multiple activators and co-activators to IE promoters, including the CBP and p300 histone acetyltransferases and the Brg1 and Brm1 components of the SWI/SNF complex (129, 175). However, none of these components were found to be necessary for viral IE gene expression (184). Interestingly, work from the Knipe lab has shown that another chromatin remodeling factor, Snf2 of ISWI, is present on IE promoters and it affects ICP0 gene expression (30). The dependence of ICP0, but not ICP4 or ICP27, expression on Snf2 suggests that different promoters may rely on different chromatin remodeling factors, perhaps based partially on the post-translational modifications of histone tails, and the cellular and viral proteins capable of binding each specific promoter (30). While specific mechanisms for chromatin remodeling on IE promoters have not been established, it is likely that RNA PolII relies on chromatin remodeling factors to translocate along the viral DNA.

The majority of recent research pertaining to chromatin remodeling during lytic infection has involved describing the mechanisms for nucleosome remodeling on IE promoters, with

significantly less attention to remodeling on E and L promoters. The Knipe lab has shown that in the absence of the IE protein, ICP0, there is an increase in H3 deposition on E promoters, and a decrease in the proportion that is acetylated (active), implicating ICP0 in this process (49). ICP0 cannot directly bind to DNA, therefore, it is likely that ICP0 acts as a global de-repressor by altering the functions of histone deacetylase complexes, and as a consequence, promoting the acetylation of histones on genomes (53, 100, 258). Acetylation can lead to binding of bromodomain containing proteins such as components of the SWI/SNF complex, perhaps suggesting a basic mechanism for chromatin remodeling on E and L promoters. However, evidence proving the presence of chromatin remodeling complexes at E or L promoters does not currently exist. Interestingly, the same study implicating ICP0 in increasing the acetylation of H3K9 on E promoters also demonstrated a moderate increase in H3 deposition on the viral genome in the absence of functional ICP4, although it was not statistically significant (49). Perhaps a combinatorial effect between ICP0, ICP4, and VP16 (all tegument proteins), leads to the activation and promoter-dependent recruitment of chromatin remodeling complexes to viral promoters.

Unlike during lytic infection, during latent infection the HSV-1 genome is associated with nucleosomes in an ordered distribution (71). Both euchromatic and heterochromatic epigenetic marks are associated with the viral genome during latency. Euchromatic marks such as acetyl- H3K9 and H3K14 as well as methyl-H3K4 are enriched in the LAT region of the genome, perhaps partially accounting for LAT production during latency (180, 181, 237). Heterochromatic marks are also found on the LAT promoter, perhaps resulting in the differential expression of LAT. Constitutive heterochromatic marks such as H3K9me3 and facultative heterochromatic marks such as H3K27me3 are found on lytic promoters on the viral genome (48,

185). None of these marks appear to be focused on a specific gene class; it is postulated that the distribution of epigenetic marks on latent genomes is highly cell specific within the neuronal population. Thus, some cells would have genomes associated with facultative heterochromatin while others with constitutive heterochromatin (21).

Reactivation of the viral genome requires the reversal of heterochromatic marks to euchromatic marks. Upon reactivation, the acetyl- H3K9 and H3K14 euchromatic marks associated with the LAT promoter become deacetylated and lytic promoters become associated with acetylated histones (4). Histone acetylation appears to play a major role in governing reactivation as treatment of latently infected cells with HDAC inhibitors results in acetylation of histones and increased viral gene expression (237). It is likely that some of the same mechanisms governing chromatin-remodeling mechanisms during lytic infection directly apply to the mechanisms of chromatin remodeling during latent infection.

Our lab has developed a model of quiescence in which a quiescent state, mimicking the latent state in vivo, is established with infection of an IE null virus (282). Our lab has shown that expression of ICP0 contributes to a hyperacetylated chromatin state, and not only helps to prevent deposition of the repressive marks H3K9me3 and HP1γ, but also promotes their removal (99, 100). The activities of ICP0 in chromatin remodeling stem not only from its enzymatic function as an E3 ubiquitin ligase, but also from its ability to interact with and disrupt the HDAC1/HDAC2/Rest/CoRest/LSD1 complex (91, 100, 101, 121). Interestingly, ICP0 does not appear to be absolutely necessary for reactivation (262). This suggests that alternative mechanisms to regulate chromatin-remodeling functions during reactivation exist. It has also been suggested that VP16 is the only protein necessary for coordinated exit from latency (287, 325), implicating VP16 contributes to chromatin remodeling during reactivation. It is likely that

ICP0's functions may be necessary for derepression of constitutive heterochromatin, whereas, perhaps other viral factors, such as VP16 or ICP4, could play a role in derepression from facultative heterochromatic states (124). Studies investigating the chromatin states of distinct pools of cells based on permissivity (LAT expression) with respect to the activities of ICP0, ICP4, and VP16 would shed valuable light onto mechanisms of chromatin remodeling during the exit from latency.

1.4.2 General mechanisms of viral transcription

Viral transcription is necessary for lytic replication and reactivation from the latent phase of the viral lifecycle. Viral transcription occurs in much the same way as cellular transcription; however, there are many notable differences. These differences enable the virus to overcome host defenses and to alter cellular transcriptional activities to promote viral transcription over cellular. Transcriptional mechanisms utilized during lytic replication and during the reactivation from latency are likely similar, with differences potentially arising from differing chromatin states of the genome and cell type specific factors. The chromatin state of the viral genome during lytic and latent infection has been addressed and thus, general mechanisms of viral gene expression and the temporal regulation of kinetic expression will be discussed.

The activation of viral gene expression relies on the architecture of the viral promoter. As mentioned previously, there are two general classes of promoters, focused and dispersed. While the majority (~70%) of cellular promoters fall into the dispersed category, all HSV promoters studied to date fall into the focused category. Additionally, while only approximately 1/3 of focused cellular promoters contain TATA boxes (~10% of total cellular promoters), all HSV-1 promoters contain TATA boxes. Additionally, HSV-1 promoters contain multiple cis-

regulatory sites for recognition by a variety of both cellular and viral factors. Cis-regulatory sites may help define the level and temporal kinetics of gene transcription.

Much like cellular genes, viral genes are transcribed by cellular RNA PolII (3), and as such use many of the same cellular transcription factors as cellular promoters. Studies have implicated the general transcription factors TFIID and TFIIA, as well as the activator, Sp1, and the coactivator, Mediator, in the formation of transcription complexes on viral promoters (162, 198, 284, 372, 373). The current understanding is that pre-initiation complex formation on viral promoters is activated by the viral factors VP16 and ICP4, leading to the recruitment of cellular RNA PolII to viral promoters *in vivo* (117, 331). The precise composition of the viral pre-initiation complexes has not been determined, although as mentioned, it is believed to contain many of the same components as cellular complexes (372).

Once complete preinitiation complexes have formed on the promoters, RNA PolII promoter release, elongation, termination, and mRNA processing and export must occur. To date, studies investigating HSV-1 mediated RNA PolII promoter release and transcription elongation have focused primarily on the phosphorylation states of the CTD of RNA PolII. It has been shown that RNA PolII is alternatively phosphorylated during HSV-1 infection, with a loss of phosphorylation on Ser2 (108, 270) and that this leads to reduced transcription of cellular genes (309). The formation of an alternatively phosphorylated form of RNA PolII is dependent on mechanisms mediated by the viral IE proteins ICP22 and ICP27 and the viral kinase UL13 (63, 108, 208). As mentioned previously, phosphorylated forms of the CTD of RNA PolII act as docking sites for chromatin remodeling complexes and mRNA processing factors. It is generally understood that Ser5 phosphorylation is important for promoter clearance while phosphorylation of Ser2 is important for proper elongation and the recruitment of specific polyadenylation and

cleavage factors (1, 172). Recall that during cellular transcription Ser2 phosphorylation generally occurs via cdk9 of pTEFb; it is speculated that because Ser2 phosphorylation is absent on RNA PolII in infected cells, pTEFb may not be involved in viral transcription, although contradictory evidence does exist (82, 269).

The loss of Ser2 phosphorylation would, in theory, increase RNA PolII pausing, and prevent efficient elongation; however, the virus has evolved mechanisms to promote efficient elongation. First, as mentioned previously, the viral genome is not associated with normally distributed histones or nucleosomes; the reduced level of histones may reduce the amount of RNA PolII pausing. Second, the viral genome is essentially intronless, making the transcribed region shorter, which may also reduce the amount of RNA PolII pausing (269). Third, it has been suggested that stalled RNA PolII, which is often present in highly transcribed regions, is specifically removed and degraded in a proteasome dependent manner, and that this effect is mediated by the phosphorylation status of the CTD of RNA PolII (63, 200). Specific elongation factors such as the SEC, Elongin SIII complex, TFIIF, or TFIIS have not been specifically implicated in promoting viral mRNA elongation.

The loss of Ser2 phosphorylation of RNA PolII also raises questions about the mechanism of recruitment of 3'mRNA processing factors. UV crosslinking studies have shown that components of the polyadenylation machinery, including the CstF, are present at the 3' end of viral mRNA transcripts (220), indicating that normal polyadenylation processes can occur in the absence of Ser2 phosphorylation. Interestingly, ICP27 was also implicated in the enhancement of polyadenylation at L, but not E genes, implying that there may be differences in the mechanisms by which polyadenylation machinery is recruited to viral templates (220). The mechanisms of recruitment of polyadenylation and cleavage factors have not yet been

investigated, although, with the lack of Ser2 phosphorylation on the CTD of RNA PolII, it is clear that HSV-1 utilizes alternative mechanisms.

As mentioned previously, viral genes are generally intronless, and therefore do not require splicing. Interestingly, however, to enhance viral transcription processes, mechanisms to inhibit cellular splicing exist. The IE protein, ICP27 is involved in promoting the aberrant phosphorylation of some SR proteins involved in spliceosome assembly. This aberrant phosphorylation pattern results in incomplete spliceosome assembly, and thus prevents the splicing of cellular transcripts (29, 290). Because splicing is intimately linked to mRNA export, the inhibition of cellular splicing in HSV-1 infected cells likely promotes the export of viral mRNA over cellular. Additionally, because viral mRNAs are mostly unspliced, and the splicing and export machinery are linked, viral mRNAs are not recognized by the nuclear export machinery in the same manner as spliced cellular mRNA. Instead, ICP27 interacts directly with the viral mRNA, the TAP-p15 complex and the cellular export factor Aly/REF to mediate viral mRNA export. If any of these interactions are abrogated, viral mRNA export does not occur (45, 46, 151, 152).

1.4.3 Regulation of the temporal kinetics of gene expression

A series of elegant studies performed in the mid to late 1970's demonstrated that, upon infection, the viral genome was transcribed in a temporal cascade with three distinct phases of gene expression, termed alpha, beta, and gamma, which are now referred to as immediate early (IE), early (E), and late (L) (135, 136, 316). IE proteins are most abundantly expressed 2-4 hpi, E proteins between 4-7 hpi, and L proteins after 7 hpi (136). IE gene expression occurs *de novo* viral protein synthesis as a consequence of a virion component, VP16, which enters the cell in

the tegument of the virus and translocates to the nucleus to activate IE gene expression (13, 38, 136). The expression of IE genes was found to be necessary for E and L gene synthesis, as evidenced by both the absence of E protein expression immediately after removal of protease inhibitor treatment at the onset of infection, and the existence of temperature sensitive mutants that were defective in E and L gene expression at the non-permissive temperature (77, 136, 251, 260, 261, 350). Many of these ts mutants have been linked to mutations within the IE transcriptional activator, ICP4 (16, 77, 261). Late gene expression has the additional requirement of viral DNA replication, although a subset of L genes that are minimally expressed in the absence of vDNA replication exist; these are referred to as γ1 or "leaky late" genes. Interestingly, concomitant with E gene expression, IE gene expression is reduced, and likewise, upon viral DNA replication, E gene expression is reduced (Figure 7). It has been established that the majority of IE genes are regulatory proteins, while E genes are associated with processes related to nucleotide metabolism and DNA replication, and L genes are primarily structural components of the virion (Reviewed in 277). The specific mechanisms that dictate the transitions between IE, E, and L gene expression remain elusive, although the chromatin state, viral promoter architectures, and the nuclear milieu of cellular and viral proteins likely play a substantial role.

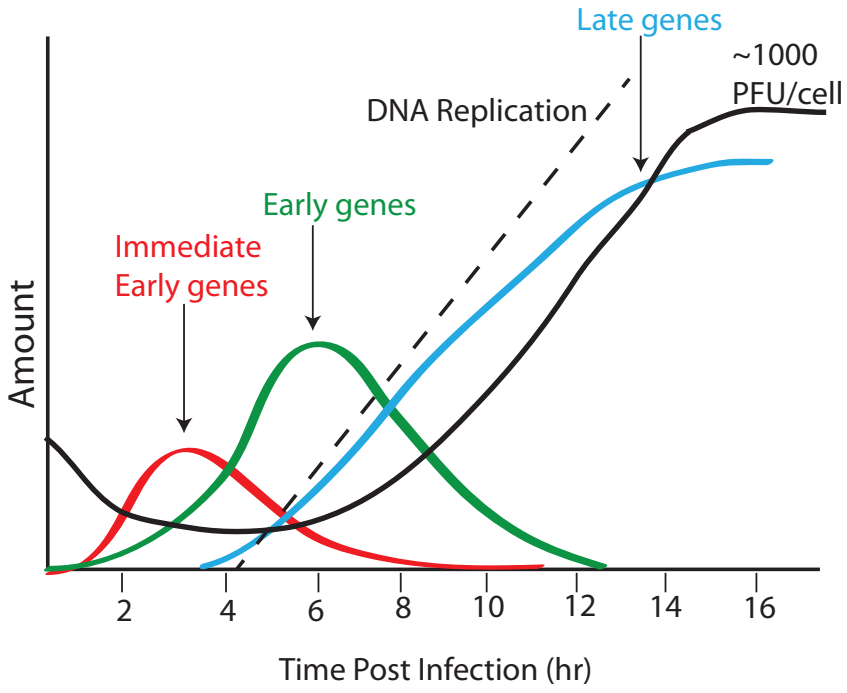


Figure 7: Temporal expression of viral genes.
Viral genes are categorized into three groups based on their requirements for expression. This graph represents the expression of viral genes in a time dependent fashion with the relative abundance of expression on the Y-axis and the hours post infection on the X-axis. This figure was adapted from (371).

Generally HSV-1 promoters corresponding to the different temporal classes (IE, E, L) are distinguishable by the complexity of activator binding sites present within the promoter (reviewed in 277). In other words, IE promoters are more complex with a higher degree of regulatory elements than E promoters. Likewise, E promoters are more complex than L promoters (Figure 8). Regulation of viral transcription arises in part from differences in promoter architecture, and the cellular factors that recognize them.

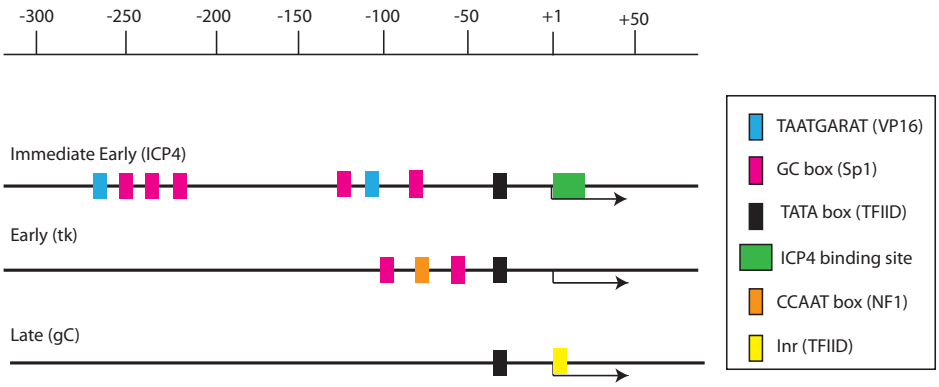


Figure 8: Promoter Architectures.
Representation of promoter architecture from each kinetic class of viral gene. Figure was adapted from (371).

IE promoters contain a binding site, with a consensus sequence of TAATGARAT, which is recognized by the IE transcriptional activator VP16. VP16 is trafficked into the cell in the tegument of the virion. Once in the cytoplasm it interacts with host cell factor (HCF), which promotes the translocation of both factors into the nucleus. In the nucleus, VP16 and HCF form complexes on IE promoters with the octamer binding protein 1 (Oct-1) to promote the transcription of IE genes *de novo* viral protein synthesis (Reviewed in 360). VP16 induces the formation of preinitiation complexes on viral DNA through interactions with multiple components of the cellular transcription machinery including TFIID, TFIIB, TFIIA, and the Mediator complex, thereby enhancing RNA PolII recruitment to IE promoters (143, 170, 204, 205, 228, 313). Additionally, VP16 mediated transcription of IE genes appears to depend on the stabilization of the VP16-Oct1-HCF complex through components of the nuclear lamina (300). In addition to VP16 mediated activation through interaction with the TAATGARAT motif, other cis-regulatory elements are present in IE promoters, including GC boxes and CCAAT binding sites, which specify for binding of Sp1 and NF1 respectively. Additional binding of these

activators likely increases the affinity and efficiency of RNA PolII binding to the promoter, thus promoting quick and efficient activation of IE genes.

Concomitant with E gene activation, IE gene expression is repressed, largely through the influence of the IE transcriptional activator, ICP4. ICP4 has been shown to bind to a strong site within its own promoter, and through the formation of a tripartite complex with TFIIB and TBP, prevents the activation of transcription (119, 120). This mechanism, however, appears to be specific for repression of ICP4 as it is dependent on a properly positioned ICP4 binding site and other IE promoters do not contain them (182). Interestingly though, in the absence of ICP4 all IE proteins are overexpressed, indicating alternative mechanisms for ICP4 induced repression exist.

Early promoters do not contain a TAATGARAT element, making them non-responsive to VP16. Instead, E genes are activated by the IE protein ICP4. ICP4 E promoter occupancy is not specifically dictated by the DNA sequence in the promoter, instead ICP4 may bind to DNA in a "less discriminant, possibly non-specific" manner (66), which will be further discussed. ICP4 functions as an activator of viral gene expression by stabilizing TFIID onto viral promoters (117). Additionally, ICP4 has been shown to form complexes with both TFIID and Mediator in infected cells (198), thus reinforcing ICP4 function as an activator of E, and L gene expression. In the absence of ICP4, E genes are not expressed, which is likely a result of inefficient recruitment of TFIID and RNA PolII to the viral promoters (284). This underscores the importance of ICP4 in preinitiation complex formation and viral gene expression. In addition to ICP4, E promoters contain binding sites for cellular activators including Sp1 and NF1 (87, 155). It is likely that Sp1, NF1, and ICP4 act together to activate E gene expression.

Just like IE genes, E genes are repressed after their initial expression. A current mechanism for the attenuation of E gene transcription does not exist, although it is likely dependent on viral DNA replication and alterations in cellular factors throughout infection. It is possible that nucleosome positioning on viral DNA is different early versus late in infection and that this could make E promoters less accessible to transcription factors. Another potential mechanism could arise from the alteration in quantity or modifications of the transcription factors required for E gene expression. It has been shown that the cellular transcription factor Sp1 becomes phosphorylated at between 4-6hpi and the phosphorylated form does not activate transcription as well *in vitro* (162). The DNA damage response protein, ATM, phosphorylates Sp1 during HSV-1 infection, potentially linking the Sp1 phosphorylation event to viral DNA replication, a prerequisite for L gene expression and E gene repression (144). Another potential mechanism for E gene repression is based on evidence from our lab that indicates TFIIA is strictly required for E gene activation, but is dispensible for L gene expression (372). Additional work suggests that TFIIA may be depleted in cells late in infection, which may lead to reduced availability of TFIIA to bind to E promoters, resulting in a decrease in E gene transcription(371).

The activation of L gene expression also relies heavily on cellular and viral activators. L genes can be divided into two sub-classes based on their requirement for DNA replication; γ1, or leaky late, genes are minimally expressed prior to viral DNA replication, and maximally expressed after, whereas the expression of γ2, or true late, genes has a strict requirement for DNA replication. Consistent with E gene expression, L gene expression relies on the viral activator ICP4. ICP4 mediated expression of L genes relies on the presence of a TATA box, which specifies for TBP binding and the Inr, which specifies for TAF1/2 binding. Our lab has reported that the Inr sequence is necessary for ICP4 activated expression of the true late gene gC

(118, 163). This is likely the result of the ability of ICP4 to function in place of TFIIA to stabilize PIC formation in an Inr dependent manner late in infection (373). While leaky late promoters do contain binding sites for cis-regulatory elements, such as Sp1, these seem to be dispensable for activation (138, 202). Instead, properly position TATA boxes and Inr elements appear to regulate L gene activation. The absence of physiologically relevant cis-regulatory elements in L promoters provides a distinction between the potential mechanisms of activation between E and L genes.

1.5 ICP4, THE MASTER TRANSCRIPTIONAL REGULATOR

ICP4 is the major transcriptional regulatory protein encoded for by HSV-1. There are two ICP4 loci because of its location within the Rs region of the genome. The presence of two ICP4 loci within the genome underscores its importance in the viral lifecycle. In the absence of ICP4, IE proteins are over produced, E and L genes are not transcribed, DNA replication does not occur, and as a consequence, new viral progeny are not formed. Many of the functional abilities of ICP4 arise from its biochemical and structural properties.

1.5.1 Biochemical and structural properties

ICP4 is 1294 amino acids and has a predicted molecular weight of approximately 133kDa. Extensive modifications of the protein, such as phosphorylation, adenylation, guanylation, and ADP-ribosylation also occur, contributing to an apparent molecular weight of approximately 175kDa, although in solution it exists as a 350kDa dimer (19, 20, 60, 95, 223,

263, 361). While ICP4 is extensively modified, only the phosphorylation of ICP4 has been shown to alter its activity by affecting the DNA binding properties of the molecule (224, 250). ICP4 has an elongated structure, with a stokes radius of about 90 angstroms (296). These studies also suggest that the N-terminus of ICP4 has a more elongated structure while the C-terminus is more globular (296). The shape of the molecule helps to define the functional abilities of the molecule; the length of the molecule in addition to its ability to multimerize on DNA increases its affinity for weak binding sites and may allow it to function at a distance (183).

The sequence of ICP4 is generally conserved amongst alphaherpesviruses. There are two large regions of conservation mapping to the middle and C-terminal regions of the molecule (see 65 for a sequence comparison). These regions correspond to a DNA binding domain, nuclear localization sequence, and a C-terminal transactivation domain. Both of these regions are functionally important for the molecule and will be discussed below. Additionally, there is a region with significantly less conservation within the N-terminus that corresponds to a N-terminal transactivation domain. Whilst the N-terminal portion of ICP4 is not highly conserved amongst alphaherpesviruses, there are small regions of conservation. Some of these regions, including a polyserine tract, are important for viral growth *in vivo* (11, 362). Additionally, while only small segments of the N-terminus are conserved, in combination with the DNA binding domain and nuclear localization sequence, it is sufficient to activate E gene expression (70).

ICP4 localizes to the nucleus of the infected cell as a consequence of the nuclear localization sequence. Immunofluorescence assays indicate that within the infected cell nuclei, there are small punctate accumulations of ICP4 that generally co-localize with other components of the transcription machinery, and some of which localize to sites of viral genome deposition (207). These punctate accumulations are referred to as pre-replication compartments. The C-

terminal transactivation domain appears to play a role in the formation of pre-replication compartments as in its absence ICP4 is evenly distributed in the nucleoplasm (207, 294). Interestingly, only a subset of the punctate accumulations of ICP4 localizes with viral DNA and expands into fully formed replication compartments. The presence of ICP4 foci that do not localize with viral DNA or expand into replication compartments has not been fully explained.

For ICP4 to function as a transcriptional activator, or repressor, it must be able to bind to DNA. Genetic studies have implicated that the regions between amino acids 300-500 approximately, are important for the DNA binding activities of ICP4, and that this region is necessary for the activation and repression of ICP4 mediated transcription (70, 253, 254, 295). Early studies suggested that ICP4 preferentially binds to DNA containing the consensus sequence ATCGTCNNNNYCGRC where N=any base, R=purine, and Y=pyrimidine, such as a site in the ICP4 promoter (93), however it was noted that other viral promoters bound by ICP4 did not contain this specific sequence. Thus, a looser consensus sequence was generated as RTCGTCNNYNYSG where S=C or G (for N, R, Y see above) (75). However, as discussed previously, while the DNA binding domain of ICP4 is required for activated transcription, the ICP4 binding sites in the viral promoters are not specifically required for ICP4 mediated activation of transcription (118). Interestingly, the ability of ICP4 to multimerize on the DNA, a function of the C-terminal transactivation domain, increases the affinity of ICP4 for viral promoters (183). Coupled with the fact that nucleosomal deposition on viral DNA during lytic infection is minimal and irregular, ICP4 may bind somewhat indiscriminately to open regions of the viral genome as a function of its DNA binding domain and multimerization ability (48, 183, 245).

ICP4 also contains two viral transactivation domains, one at the N-terminus and one at the C-terminus. Aside from conferring for multimerization, the C-terminal transactivation domain interacts with components of the cellular transcription machinery including TAF1 of TFIID (41). The C-terminus of ICP4 is necessary for viral DNA replication and L gene transcription, although the precise reasons are unknown. Defects in DNA replications may result from reduced E gene expression in C-terminal truncation mutants. As E genes are primarily responsible for nucleotide metabolism and DNA replication, it is possible that E proteins are not produced to a sufficient quantity to drive viral DNA replication, which is a prerequisite for L gene transcription. L gene expression may also result the multimerization function of ICP4, which acts to increase the affinity of ICP4 to L promoters, which, recall, do not have cellular activator binding sites. In fact, ChIP assays indicate that the C-terminus is required for ICP4 binding on L promoters (284). Additionally, as mentioned previously, L promoters contain Inr sequences in addition to TATA boxes; it is possible that the C-terminus of ICP4 may be important for forming complexes on promoters containing Inr sequences, since TAF1 recognizes Inr and the C-terminus of ICP4 interacts with TAF1. Evidence herein and elsewhere suggest that there are redundancies within the individual transactivation domains of ICP4, and interaction with TFIID is one of them, implicating that even in the absence of the C-terminus, ICP4 may be capable of interacting with TFIID.

The N-terminal transactivation domain, as mentioned previously, is less conserved amongst alphaherpesviruses. As mentioned above, the N-terminal transactivation and DNA binding domains of ICP4 are sufficient to activate E gene transcription. Large deletions within the N-terminus appear to have moderate effects on E and L gene transcription in cell culture (11, 295, 362). The N-terminus has been implicated in the repression functions of the molecule. In

vitro studies implicate that the N-terminal transactivation domain of ICP4 is important for forming a tripartite complex with TFIIB and TBP at strong ICP4 binding sites, and that this complex prevents the activation of viral gene transcription (119, 182, 303). Thus, both the N-terminus and C-terminus have a role in interacting with TFIID through TBP and TAF1 respectively. Additionally, regions of the N-terminus of ICP4 have been shown to be important for viral growth within the trigeminal ganglia, and as a result, reactivation. Interestingly, one of the regions important for viral growth *in vivo* is the polyserine tract, which is heavily phosphorylated by PKA (11, 361, 362). As mentioned previously, differences in phosphorylated forms of ICP4 have been shown to have alterations in DNA binding affinity m(224). While mutants lacking the polyserine tract do show a reduced ability to bind to strong ICP4 binding sites, as discussed, the ability of ICP4 to bind to DNA is likely indiscriminant, and relies partially on the multimerization properties of the molecule (340). It is instead, more likely, that regions within the N-terminus of ICP4 provide a function that is necessary in neuronal cells and either unnecessary or redundant in epithelial cells. Whether this function(s) is related to the differing chromatin states of the viral genome, or the cell type specific activators in the neuron has not been determined, and remains a focus of our lab.

1.5.2 Mechanisms of action

ICP4 is necessary for both activation and repression of viral gene transcription. ICP4 autoregulation requires a properly positioned ICP4 binding site, the DNA binding capabilities, and the formation of a tripartite complex on DNA. Evidence has shown that the strength of the ICP4 binding site, and its relative position to the TATA box are important in defining autoregulation (119, 182, 197). Interestingly, ICP4 binds to the major groove, and TBP binds to

the minor groove of DNA, and optimal repression occurs when their binding sites are ½ a helical turn apart (74, 197). This is important because it was previously believed that the mechanism for repression was by physically blocking the promoter region, making it inaccessible to the transcription machinery (233). Instead, this data argues that an alternative mechanism for ICP4 mediated repression exists. It was determined in the mid-1990's that ICP4 forms a complex with TFIIB and TBP on DNA and that this complex was important for repression. Mutants capable of binding to DNA, but incapable of forming this complex were incapable of the repression of transcription. Additionally, evidence suggests that the formation of this complex prevents activator mediated transcriptional activation (119, 120, 182).

Interestingly, there are only three strong ICP4 binding sites within the genome that are associated with repression. They are located in the promoter regions of ICP4, LAT, and OrfP. The presence of strong binding sites within these promoters suggests that in the presence of ICP4, these transcripts would not be expressed, however in the absence of ICP4 they may be greatly expressed (reviewed in 66). LAT is highly expressed during latency, when ICP4 is not produced. Additionally, miRNAs targeted to ICP0, ICP4, and γ34.5 reside within the OrfP region of the genome, implicating that ICP4 may be capable of repressing the transcription of these miRNAs upon reactivation, perhaps leading to a more robust activation of viral gene expression (334). These data suggest a potential role for ICP4 in the maintenance of or reactivation from latency. The precise role of ICP4 in reactivation from latency and the significance of miRNAs in this process are not understood to date.

The mechanisms related to the activation of gene transcription suggest that ICP4 binds to viral DNA (discussed above) and stabilizes TFIID on either the TATA box or Inr sequence (163, 373). The formation of TFIID-ICP4 complexes at E viral promoters is partially dependent on the

TATA box, suggesting that ICP4 and TFIID cooperate in binding at viral promoters (284). Given the indiscriminant manner with which ICP4 binds to DNA, TFIID specific binding sites and potentially histone deposition may help position ICP4 at the promoter region. TFIID nucleation of the core promoter is a pivotal step for pre-initiation complex for cellular transcription, which is likely also true for viral transcription. It is probable that preinitation complexes form on viral promoters in much the same way as on cellular promoters, as most of the same GTFs are also necessary for viral transcription (41, 372). It has also been recently demonstrated that ICP4 can form complexes with TFIID and Mediator in infected cells, indicating the importance not only of the general transcription factors, but co-activators as well (198). This is likely a function of the large size and elongated structure of ICP4 providing multiple surfaces for transcription factor, activator, and coactivator binding. To date, ICP4 has not been implicated in other transcriptional processes outside of preinitiation complex formation. It is evident that ICP4 can interact with a variety of cellular transcription factors, but the molecular basis for many of the aforementioned functions and interactions have not yet been defined.

1.6 RATIONALE

Herpes Simplex Virus is a ubiquitous pathogen with detrimental effects on immunocompromised hosts. Infection of immunocompromised hosts, generally neonates, transplant patients, and HIV infected individuals, can lead to a disseminated infection, encephalitis, and in some cases death. Infection with HSV also leads to increased rates of HIV transmission and ocular infection can result in blindness. Currently acyclovir and derivatives are available for treatment of HSV,

however drug resistance poses a problem, particularly in the immunocompromised. Other drugs are available for HSV treatment but they are generally toxic and the delivery routes are not ideal. Thus, the development of new nontoxic drugs or a vaccine would be beneficial for public health. Additionally, the potential for HSV as a therapeutic agent is under investigation. The development of novel drugs, vaccine strategies, and therapeutic agents requires an understanding of the basic molecular virology of HSV.

For efficient viral replication, transcription of the viral DNA genome must occur. The IE protein, ICP4, is required for transcription of approximately 80 of 85 viral genes, underscoring its importance in viral transcription and replication. To activate viral transcription ICP4 forms complexes with components of the cellular transcription machinery, such as TFIID and Mediator (198). Investigating the structure of ICP4, defining regions that are important for interactions with the cellular transcription machinery, and identifying novel transcription complexes with which ICP4 interacts can lead to advancements in the development of novel treatments. The proof of principle for this exists in a study in which transgenic mice were created that contained a mutant form of ICP4 that reduced the activity of wtICP4 in cell culture. The transgenic mice were challenged with wild type virus and the antiviral activity of the mutant ICP4 was assessed. The study demonstrated reduced viral replication both in the eye and the trigeminal ganglia in the transgenic mice (304). Importantly, this study demonstrated that altering the activity of ICP4, even moderately, could greatly affect the ability of the virus to replicate *in vivo*. Therefore, the development of a drug that would disrupt ICP4 activities, such as its interactions with TFIID, would prevent activated transcription from the viral genome and consequently reactivation.

Therefore, the goals of these studies were to more accurately define the molecular basis for ICP4 mediated functions with respect to both the structure of the molecule and the transcriptional complexes with which it interacts. More specifically, the goals of this project were:

i. **To determine the contribution of conserved regions within the N-terminus to viral transcription, both in the presence and absence of the C-terminal activation domain**
ii. **To investigate ICP4-mediated protein interactions throughout infection.**

These goals will be addressed in the remaining chapters of this thesis.

2.0 THE N-TERMINUS AND C-TERMINUS OF HSV-1 ICP4 COOPERATE TO ACTIVATE VIRAL GENE EXPRESSION

2.1 ABSTRACT

ICP4 activates transcription from most viral promoters. Two transactivation domains, one N-terminal and one C-terminal, are largely responsible for the activation functions of ICP4. A mutant ICP4 molecule lacking the C-terminal activation domain (n208) efficiently activates many early genes, while late genes are poorly activated, and virus growth is severely impaired. The regions within the N-terminus of ICP4 (amino acids 1-210) that contribute to activation were investigated by analysis of deletion mutants in the presence and absence of the C-terminal activation domain. The mutants were assessed for their abilities to support viral replication and to regulate gene expression. Several deletions in regions conserved in other α-herpesviruses resulted in impaired activation and viral growth, without affecting DNA binding. The single small deletion that had the greatest effect on activation in the absence of the C-terminus corresponded to a highly conserved stretch of amino acids between 81-96, rendering the molecule nonfunctional. However, when the C-terminus was present, the same deletion had a

minimal effect on activity. The amino terminus of ICP4 was predicted to be relatively disordered compared to the DNA binding domain and the C-terminal 500 amino acids. Moreover, the amino terminus appears to be in a relatively extended conformation as determined by the hydrodynamic properties of several mutants. The data support a model where the amino terminus is an extended and possibly flexible region of the protein, allowing it to efficiently interact with multiple transcription factors at a distance from where it is bound to DNA, thereby enabling ICP4 to function as general activator of polII transcription. The C-terminus of ICP4 can compensate for some of the mutations in the N-terminus, suggesting that it either specifies redundant interactions or enables the amino terminus to function more efficiently.

2.2 INTRODUCTION

The Herpes Simplex Virus Type 1 (HSV-1) genome is transcribed in a highly regulated, sequential cascade (135, 136) by the cellular RNA polymerase II machinery (3). VP16 is a potent transactivator that is carried into the cell in the tegument of the virion (13) and activates the transcription of the Immediate Early (IE) genes (13, 38, 259). Upon subsequent protein synthesis, the IE protein ICP4 acts as a transcriptional activator, promoting the expression of Early (E) and subsequently Late (L) genes (77, 114, 260, 261, 349, 350). In addition to the activation function, ICP4 can also be a repressor of transcription in some contexts (68, 120, 244, 273). As a consequence, temperature sensitive (*ts*) and deletion mutants of ICP4 over express IE proteins, but are highly defective for E and L gene expression (67, 77, 261, 350).

ICP4 is a 175 kDa nuclear phosphoprotein (60, 255) that exists as a dimer in an elongated conformation (223, 296). ICP4 is also a sequence specific DNA binding protein (75, 94). The

binding to specific sites is required for the repression of the ICP4, LAT, and OrfP/L/ST promoters (97, 119, 188, 233, 273). In contrast, while the DNA binding activity of ICP4 is required for activation of transcription, individual specific binding sites are not (52, 87, 90, 125, 302).

Alignment of the amino acid sequences from α-herpesvirus ICP4 homologues reveals two large blocks that are highly conserved amongst the homologues between amino acids 300-500 and the C-terminal 500 amino acids, with respect to HSV-1 ICP4. The extensive homology shared between these homologues is suggestive of the importance of these regions (219). These regions have been shown to specify a DNA binding domain (253, 295) and a region important for transactivation, respectively (69, 70, 253, 295). There is minimal conservation among ICP4 homologues within the N-terminus of the molecule. This apparent lack of homology, however, is not indicative of importance of this region. This region has been shown to be important for the ability to activate (296) and repress transcription (119). While the N-terminus of ICP4 shares little homology amongst α-herpesviruses, a sequence alignment of the N-terminus of ICP4 homologues from HSV-1, HSV-2, and herpes B virus, reveals multiple conserved blocks of amino acids between the three homologues (Fig.9). These conserved regions may confer many of the functions of ICP4 that are attributed to the N-terminus, such as the ability to form transcription complexes on early promoters (117, 284) and the ability to repress transcription in a promoter dependent manner (119). It is also possible that some of these conserved blocks of amino acids function in conjunction with the C-terminus, perhaps forming interfaces with many of the same transcription factors.

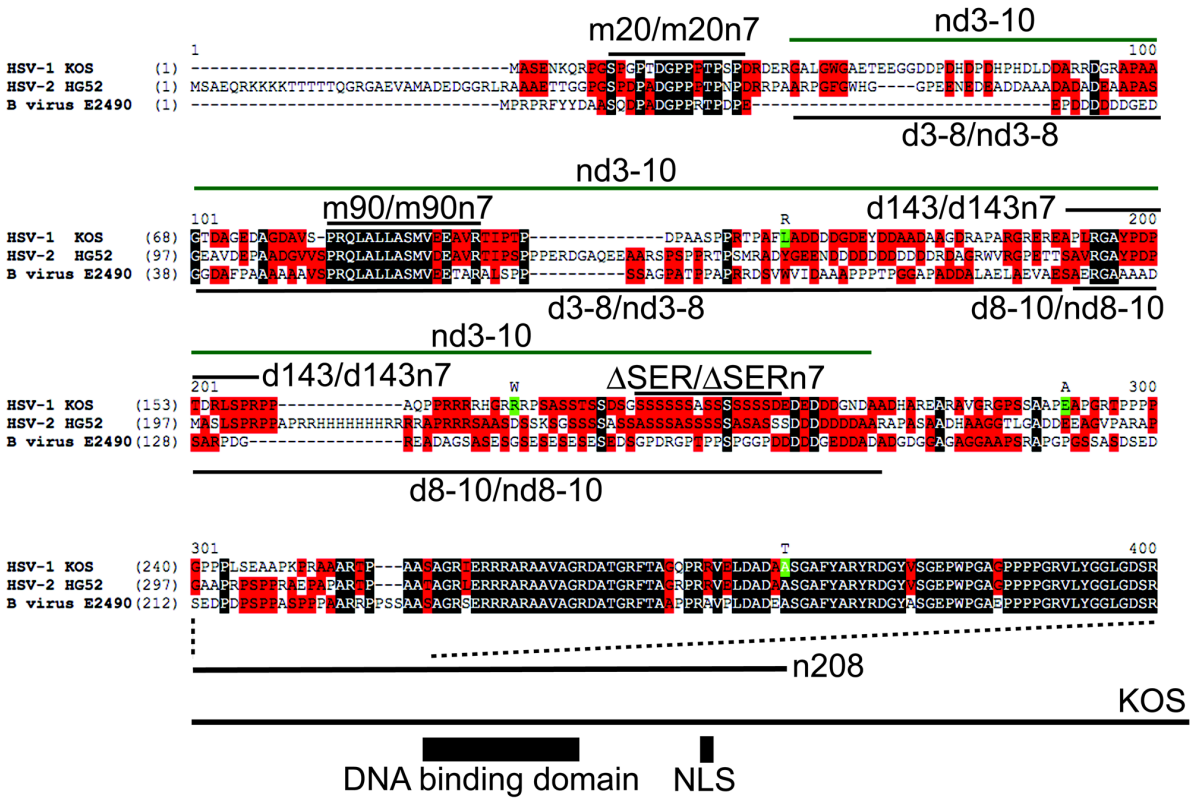


Figure 9. Sequence alignment of the N-terminus of ICP4 from HSV-1 strain KOS, HSV-2, and Herpes B Virus.

The predicted translation products of HSV-1 strain KOS, HSV-2 strain HG42, herpes B virus strain E2490 ICP4 molecules were aligned using Vector NTI. Amino acids highlighted in black are highly conserved among all three viruses while those highlighted in red are only conserved amongst two of the viruses. The black lines above and below the alignment indicate the deletions constructed within the N-terminus of ICP4. The deletions were constructed in both the wild type KOS background (before the backslash) and the C-terminal deletion mutant n208 (after the backslash).

The N- and C-terminal transactivation domains are necessary to varying degrees for the transactivation and repression functions of ICP4. For example, the C-terminal activation domain of ICP4 is necessary for efficient L gene activation, but not E gene activation (69). The N-terminal activation domain contains the site-specific repression functions of the molecule and some activation functions (119, 296). Additionally, it appears that the different regions of ICP4 have differing roles during the lytic cycle in epithelial cells as opposed to in neuronal cells (11, 362). Deletion of the conserved polyserine tract within the N-terminus of ICP4 results in a

mutant that has minor defects in replication in Vero cells and at peripheral sites of ocular infection in mice, but does not replicate within the trigeminal ganglia of mice (11, 362). Thus various regulatory functions of ICP4 are specified by different parts of the protein, most likely as a result of interactions with different components of the host transcriptional machinery, which in turn may be of varying importance in different cell types or at different times post infection.

In order to more accurately define regions of importance, their functions, and whether the C-terminus can contribute to these functions, several deletion mutants within the N-terminus of ICP4 were constructed both in the presence and absence of the C-terminal activation domain. Mutants were utilized in transient assays and assessed for their ability to complement an ICP4 null virus. Additionally, the mutations were constructed within both ICP4 loci of the viral genome. The viral mutants were analyzed for their ability to replicate, bind to DNA, and activate gene expression. The results indicate that there are multiple regions of importance within the N-terminus of ICP4, and that the C-terminus of ICP4 may cooperate with the N-terminus to regulate viral gene expression.

2.3 MATERIALS AND METHODS

Cells and Viruses. Vero and E5 cells were maintained as previously described (67). E5 cells express complementing levels of ICP4 (69). Viral mutants were constructed within the background of HSV-1, strain KOS. The ICP4 mutants n208(70), d8-10 (362), nd8-10 (283), ΔSER(11), d120 (67), and n12 (70) have been previously described. ΔSERn7 was constructed by coinfecting E5 cells with n208 and ΔSER, and screening progeny plaque isolates by PCR and southern blot hybridization for both copies of the n208 and ΔSER alleles. The mutants d3-8,

nd3-8, m20, m20n7, m90, m90n7, d143, d143n7, and nd3-10 were constructed by marker transfer (67). E5 cells were cotransfected with 2μg of the mutant plasmid DNA digested with EcoRI and 2μg viral DNA from either KOS, n12, or n208 using Lipofectamine 2000 (Invitrogen) according to the manufacturers suggestions. Plaque isolates of the transfection progeny were then screened by PCR and southern blot hybridization for both copies of the intended alleles. Mutant viruses were plaque purified a minimum of three times.

Recombinant plasmids. The plasmids pK1-2 , pn7, d3-8, nd3-8, d8-10, and nd8-10 have been previously described (69, 295). pK1-2 and pn7 encode the promoter and gene for wt (strain KOS) and n208 ICP4, respectively. pK1-2 and pn7 were transformed into E. coli GS1783 (gift of Greg Smith, Northwestern University) and site-directed deletions were constructed using the RecET recombination system as previously described (327, 328). The primers used for the construction of the corresponding mutants were: m20, CCCGCATCGGCGATGGCGTCGGAGAACAAGCAGCGCCCCGGC**CTGCAG**GACCGCG ACGAGCGGGGGTAGGGATAACAGGGTAATCGATTT and CGTCTCCGCGCCCCACCCGAGGGCCCCCCGCTCGTCGCGGTC**CTGCAG**GCCGGGG CGCTGCTTGTTGCCAGTGTTACAACCAATTAACC; m90, GCGGGCACCGACGCCGGCGAGGACACCGGGGACGCCGTCTCG**CTGCAG**ACGATCC CGACGCCCGACTAGGGATAACAGGGTAATCGATTT and GGTCCGGGGCGGCGAGGCCGCGGGGTCGGGCGTCGGGATCGT**CTCCAG**CGAGACG GCGTCCCCGGTGCCAGTGTTACAACCAATTAACC; d143, GCCGGCGACCGGGCCCCGGCCCGGGGCCGCGAACGGGAGGCC**CTGCAG**CCGCCG GCCCAGCCGCCGTAGGGATAACAGGGTAATCGATTT and

CCACCGCCCGTGACGACGTCTCCGCGGCGGCTGGGCCGGCGG**CTGCAG**GGCCTCC CGTTCGCGGCCGCCAGTGTTACAACCAATTAACC. The six nucleotides in bold correspond to PstI sites, which were added at the site of the deletion for diagnostic purposes. The primers were synthesized and gel purified by IDT (Coralville, Iowa). The kanamycin resistance gene from pEPKan-S (328) was amplified by PCR using the Failsafe PCR kit from Epicenter Biotechnologies with the following conditions: 50 ng pEPKan-S, 2uM primers, and 0.5U Failsafe Enzyme were combined in a 1X buffer provided by the manufacturer. The reactions were cycled at 96° for 5minutes, then 35 times at 96° for 30 seconds, 60° for 1 minute, and 72° for 1.5 minutes, and finally at 72° for 5 minutes. The PCR product was gel purified and transformed into GS1783 bacterial cells containing the ICP4 plasmids pK1-2 and pn7 as described (328). Transformants containing the appropriate mutant plasmids were selected for by resistance to kanamycin and carbanecillin and screened by restriction endonuclease and agarose gel analysis. Isolates containing the appropriate insert that lacked the wt allele were grown at 30° for 30 minutes in Luria Broth containing 1% arabinose. The cells were transferred to 42° for 15 min to induce the recombination system and concomitant removal of the kanamycin cassette. The cells were allowed to replicate for another 2 hours at 30° and plated on agar plates containing 100 ug/mL carbanecillin and 1% arabinose. Colonies were screened for their sensitivity to kanamycin and possession of the appropriate allele by restriction endonuclease and agarose gel analysis.

Transient complementation assay. 5 x 10^5 Vero cells in 35 mm petri dishes were transfected with 2 μg of plasmid DNA using 5μL Lipofectamine 2000 as described by the manufacturer (Invitrogen). Twenty-four hours later the cells were infected with the ICP4 mutant virus d120 at

a MOI of 1 PFU/ml. Twenty-four hours post infection the cells were scraped into the medium, disrupted by 3 freeze-thaw cycles, sonicated, and the viral lysates were clarified by centrifugation at 3000 rpm for 10 minutes. Total viral yield was determined by plaque assay on E5 cells.

Single Step Growth Analysis. 5 x 10^5 Vero cells in 35 mm petri dishes were infected at a MOI of 5 PFU/cell in 0.1 ml TBS (137 mM NaCl, 5 mM KCl, 0.5 mM MgCl, 0.68 mM $CaCl_2$, 25 mM Tricine; pH 7.35), at room temperature for 1hour. After viral adsorption, the cells were washed three times in TBS to remove un-adsorbed virus and 37º medium was added. The infection was allowed to progress for 2, 4, 8, 12, 24, and 36 hours at 37°. At these times, the infected cells were scraped into the medium, disrupted by 3 freeze-thaw cycles, sonicated, and viral lysates were clarified via centrifugation at 3000 rpm for 10 minutes. Total viral yield was determined by plaque assay on the E5 cells.

Analysis of Viral Proteins. Viral proteins were analyzed by metabolic labeling and sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and by western blot analysis. 5 x 10^5 Vero cells in 35 mm petri dishes were infected at a MOI of 10 PFU/cell in 0.1 ml TBS at room temperature for 1hr, after which time the inoculum was removed, 37º medium was added, and cells were incubated at 37° for the indicated times. The infected cells were washed in cold TBS containing 0.1mM *N-p*-tosyl-L-lysine chloromethyl ketone (TLCK) and scraped into 200µL sample buffer (0.05M Tris-HCl, pH 7, 2% SDS, 0.01% bromophenol blue, 5% sucrose, 5% β-mercaptoethanol). The infected-cell extracts were denatured at 95°C for 5 minutes and separated by SDS polyacrylamide gel electrophoresis (SDS-PAGE). Polypeptides were separated through

a 4% polyacrylamide stacking gel (4% acrylamide, 0.19% DATD, 119mM Tris, 0.1%SDS; pH7.0) and a 9% polyacrylamide separating gel (9% acrylamide, 0.4% DATD, 375mM Tris, 0.1% SDS; pH 8.8) in Tris-Glycine Electrode buffer (25 mM Tris, 192 mM glycine, 0.1% SDS; pH 8.5) at 100 V.

For analysis of metabolically labeled viral polypeptides, the medium (methionine and cysteine free DMEM) on the infected cells was removed at 3.5, 7.5, and 11.5 h post infection and replaced with 1 ml 37º TBS containing 25 μCi ^{35}S-methionone (>1000 Ci/mmole, MP Biomedicals). The cultures were incubated for an additional 30 min at 37º, after which time the labeling medium was removed and the monolayers were washed with cold TBS containing 0.1 mM TLCK, solubilized, and analyzed by SDS-PAGE as described above. Following electrophoresis the gel was fixed in H_20:methanol:acetic acid (6:3:1), dried and exposed to X-ray film (HyBlot CL from Denville).

For western blot analysis, the SDS-PAGE separated polypeptides were transferred from the polyacrylamide gel to nitrocellulose membranes by electroblotting in 1X transfer buffer (25 mM Tris, 192 mM Glycine, 20% Methanol) for 1.5 h at 250 mA. The membrane was subsequently blocked by incubation at 4°C overnight in 5% dry milk in TBS (50 mM Tris, 150 mM NaCl; pH 7.5) and then incubated with ICP4 specific antibodies (N15 or 58S) at a 1:500 dilution in TBS + 0.05% Tween 20 (TBS-T) + 1% milk for 1 hour. The immunoblot was then washed in 1% milk in TBS-T, 4 times for a total of 30 minutes. It was then incubated with either goat anti mouse (for 58S) or donkey anti rabbit (for N15) IRDye-conjugated (LI-COR) secondary antibodies at a 1:10,000 dilution in TBS-T, subsequently washed 4 times for a total of 30 minutes in TBS-T, and finally twice for a total of 2 minutes in TBS. Images were obtained using the LI-COR Odyssey Infrared Imager.

Southern Blot hybridization. DNA was isolated from infected cells and analyzed by Southern blot hybridization as previously described (282, 307). Infected cell DNA was digested with the restriction endonucleases BamHI and PstI and separated by electrophoresis in a 1.2% agarose gel. A plasmid containing the BamY fragment of the HSV-1 genome was used as the probe DNA and was labeled with ^{32}P-CTP and ^{32}P-GTP (GE Healthcare) using a Nick Translation System (Invitrogen) according to the manufacturers suggestion. The results were visualized by autoradiography.

Northern blot analysis. 5 x 10^6 Vero cells in 100 mm petri dishes were infected at an MOI of 10 PFU/cell at room temperature for 1 hour. The viral inoculum was removed, 37º medium was added, and cells were incubated at 37° for 4 and 8 hours. The monolayers were then rinsed once with cold TBS, and 2 mL Trizol reagent (Invitrogen) was added. The infected cells were scraped into the Trizol reagent and RNA was harvested according to the manufacturers (Invitrogen) guidelines. RNA was treated with 10 U RNAse free DNAse I at 37° for 15 minutes, and ethanol precipitated.

Total RNA was quantified using the NanoDrop 2000 (Thermo-Fisher). RNA samples were prepared for electrophoresis by diluting 13 μg RNA into 4.5 μL water and adding 2 μL 10X MOPS running buffer (1X MOPS running buffer; 20 mM MOPS (3(-N-morpholino)propanesulfonic acid), 1 mM EDTA, 5 mM sodium acetate), 3.5 μL 37% formaldehyde, and 10 μL formamide followed by heating the samples for 15 min at 55°. Two microliters of loading buffer (50% glycerol, 1 mM EDTA, 0.4% Bromophenol blue, 0.4% Xylene Cyanol) was added to each sample and then the samples were run on a 1.3% agarose-

formaledehyde gel (1.3% agarose, 8% formaldehyde in 1X MOPS running buffer) at 100 V for 4 hours in re-circulating 1X MOPS running buffer (20 mM MOPS, 1 mM EDTA, 5 mM sodium acetate). The gels were prepared for transfer by washing them for 45 min each in; water, 50 mM NaOH 10 mM NaCl, 100 mM Tris-HCl pH 7.5, and finally 10X SSC (1.5 M NaCl, 1.5 M Sodium Citrate). The RNA was transferred to Nytran N (Whatman) overnight in 10X SSC and the RNA was crosslinked to the membrane using a UV-Stratalinker (Stratagene). Membranes were probed with ^{32}P-nick translated DNA specific to ICP4 (BamY fragment), ICP0 (plasmid pW3ΔHS8 containing the ICP0 coding sequence), tk (the SacI-Sal I fragment corresponding to +555 to +1217 relative to the start site), LAT (+5557 to +7559 of the BamB fragment), or gC (plasmid pSXgC containing the gC coding sequence) as previously described (140). Results were visualized via autoradiography.

DNA binding assays. Protein extracts were prepared by scraping 10^6 infected (MOI = 10 PFU/cell for 6 h) Vero cells in 60 mm petri dishes into TBS + 0.1 mM TLCK, and pelleting the cells at 12.5 K for 20 sec. The pellet was then resuspended in 30 μL TE (Tris-HCl pH 8.0, 1 mM EDTA), and 30 μL 2X Lysis Buffer (100 mM Tris; pH 8.0, 1 M KCl, 4% NP-40) and 0.1 mM TLCK. The samples were incubated on ice for 45 minutes. Cell debris was pelleted at 12.5 K for 10 minutes and the supernatant was used for the binding reaction. Binding reactions were prepared with 3 μL of the protein extract, 1 ng of ^{32}P end-labeled probe DNA, and 2 ug the nonspecific inhibitor dI•dC (polydeoxyinosinate-polydeoxycytidylate) (Midland Certified Reagent Company) in binding buffer (10 mM Tris pH=7.5, 1 mM EDTA, 5% glycerol, 0.1% NP40, 50 mM KCl, 0.1 mM TLCK). The probe DNA utilized in this assay was a fragment of the ICP4 promoter termed P4, which covers -108 to +27 bp relative to the transcription initiation

site of ICP4. The binding reactions were incubated at room temperature for 30 min and then run on a 4% polyacrylamide gel in 0.5 x TBE (1 X TBE; 88 mM Tris, 88 mM boric acid, 1.25 mM EDTA) at 200 V for 1 hr 45 min. The gel was then dried and the results were visualized by autoradiography.

Size exclusion chromatography. Infected cell extracts were prepared as described above for the DNA binding assay. The extracts were run on calibrated superpose 6 10/300 column (GE Healthcare) and the fractions analyzed for ICP4 as previously described (293, 296). The LMW and HMW gel filtration calibration kits (GE Healthcare) were used to calibrate the columns. The Stokes radii were determined as previously described (293, 296) and shown in Figure 5.

Immunofluorescence. 2 x 10^5 Vero cells were seeded onto glass coverslips in a 12-well plate. Cells were infected at a MOI of 10 PFU/cell for 1 hour with the indicated viruses at room temperature. The inoculum was replaced with 37º media and the infections were allowed to progress at 37° for 2 hours. The monolayers were then fixed for 10 minutes in a 4% paraformaldehyde solution, and subsequently washed 3 times for 5 minutes each in PBS (137 mM NaCl, 2.7 mM KCl, 5.3 mM Na_2HPO_4, 1.7 mM KH_2PO_4; pH 7.4), followed by permeablization in PBS + 0.4% TritonX100 + 1% BSA for 30 min. The permeablized cells were incubated with a primary antibody for ICP4 (N15; rabbit polyclonal serum) at a 1:500 dilution in PBS + 1%BSA. The cells were then washed in PBS +1% BSA three times for a total of 30 min and twice in PBS for 20 min before being incubated in Alexa-Fluor conjugated secondary antibodies (Santa Cruz) at a 1:500 dilution in PBS. The cells were then washed 6 times for a total of 1hr in PBS and the coverslips were mounted onto a glass slide using Immumount

(Electron Microscopy Services). Images were obtained using the Olympus Fluoview FV1000 confocal microscope.

2.4 RESULTS

A mutant ICP4 molecule expressed from the virus n208 lacks the C-terminal activation 520 amino acids due to truncation. The n208 molecule retains the ability to repress transcription from the ICP4 promoter, efficiently activates early promoters, but true late gene expression is substantially impaired, resulting in relatively poor viral growth (70). Deletion of amino acids 30-274 from the n208 molecule results in a dimeric protein that can bind to DNA, yet it cannot activate or repress transcription (118, 119, 296), suggesting a role for the N-terminal region in both repression and activation. While this region may sufficiently specify interactions involved in repression and activation, it is also possible is that N-terminal activation domain and the C-terminal activation domain function together to constitute the crucial functions of ICP4. To investigate the functions of the N-terminal and C-terminal activation domains with respect to the regulatory functions of ICP4, deletions within the N-terminus of ICP4 were constructed in the context of the entire 1294 amino acid protein and in the background of the truncated n208 molecule.

Sequence alignments of ICP4 homologues between multiple members of α-herpesviridae showed very little homology within the N-terminus of ICP4. However, a sequence alignment between the primate α-herpesviruses, HSV-1, HSV-2, and herpes B virus revealed several relatively conserved regions within the N-terminus of ICP4 (Fig.9). The sequence shown for HSV-1 is the N-terminus of the KOS strain. Differences between the KOS amino acid sequence

and strain 17 (219) are highlighted in green with the strain 17 residue given above the main sequence. The sequences highlighted in black are conserved in all three viruses, while those highlighted in red are conserved in two out of three.

Deletions were constructed as indicated in Fig. 9 such that a PstI site was also inserted into the site of the deletion. The nomenclature for the newly constructed mutants, m90/m90n7 for example, represents the indicated deletion in the full-length molecule, and the C-terminal truncation mutant n208 (n7). The Δser mutants were previously constructed (11). The d3-8 and d8-10 mutants were also previously constructed (295), with the nd version representing the deletion within the n208 molecule. The resulting clones containing the deletions of interest, were verified by restriction digestion and agarose gel electrophoresis.

The mutant plasmids were assessed for their ability to complement an ICP4 null virus, d120, as any defects in complementation ability may be predictive of defects in viral transcription. Briefly, ICP4 mutant plasmids were transfected into Vero cells. The transfected cells were used as hosts for infection with d120. Transfected wt ICP4 expressed from the plasmid pK1-2 resulted in an approximately 10^5– fold increase in the yield of d120 over mock-transfected cells (not shown). Figure 10 displays the results of the complementation assay for each mutant as a percent of the yield resulting from the pK1-2 transfection. Of the mutants containing the C-terminal 500 amino acids, m90, d8-10 and d3-8 were partially impaired while the mutant that deleted the whole region d3-10 was highly impaired. Truncation of the C-terminus (n208) resulted in a 300-fold decrease in the yield of d120. The deletions within the N-terminus of the n208 molecule had a similar effect to the deletions within the intact molecule, with m90n7, nd3-8, nd8-10 and nd3-10 yielding very little to no d120 in excess of mock transfections. These defects in complementation ability are most likely reflective of reduced transcription of essential

viral genes and predictive of the possible phenotypes of viruses baring these mutations. In addition, transient transfections of the plasmids followed by western blot analysis revealed that each mutant plasmid specified a protein of the approximate expected molecular weight demonstrating that defects in complementation were not the result of stability issues or unintended gross rearrangements (data not shown).

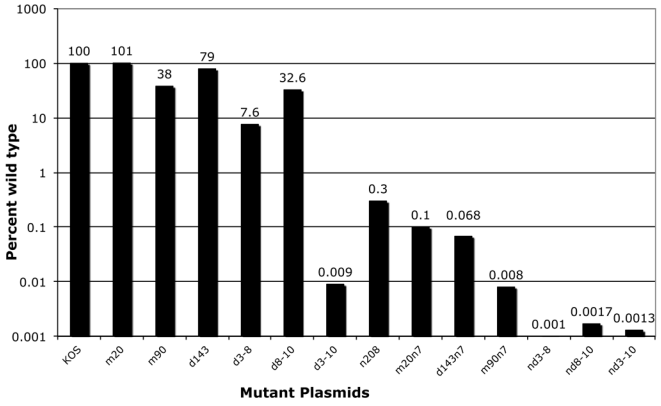


Figure 10. Complementation activity of mutant plasmids.
ICP4 mutant expressing plasmids were transfected into Vero cells and their ability to complement the ICP4 null virus, d120, was determined as described in the materials and methods. The activity is expressed as the percent yield relative to the transfection with the plasmid expressing wt ICP4 (pK1-2).

To more thoroughly assess the effects of the mutations on the viral life cycle, the deletions were constructed in both ICP4 alleles within the wild type and n208 genomes, as described in the materials and methods. The d3-10 construct proved to be transdominant-negative and therefore we were unable to construct a virus with this mutation. The PstI insertion at the deletion in all the mutants was used as a diagnostic for the intended mutations. Mutant viral DNA was digested with BamHI and PstI and separated via agarose gel electrophoresis. Southern blot analysis was performed using the BamY fragment of the genome as a probe. Figure 11 shows the results from the Southern blot analyses of the viral deletion mutants. As expected, the wild type and n208 genomes, which do not have a PstI site, yielded a fragment of 1.84 kb. PstI cleaved this 1.84kb

fragment from the deletion mutants into two smaller fragments (Figure 11), indicating that the genotypes of the mutant viruses were as intended.

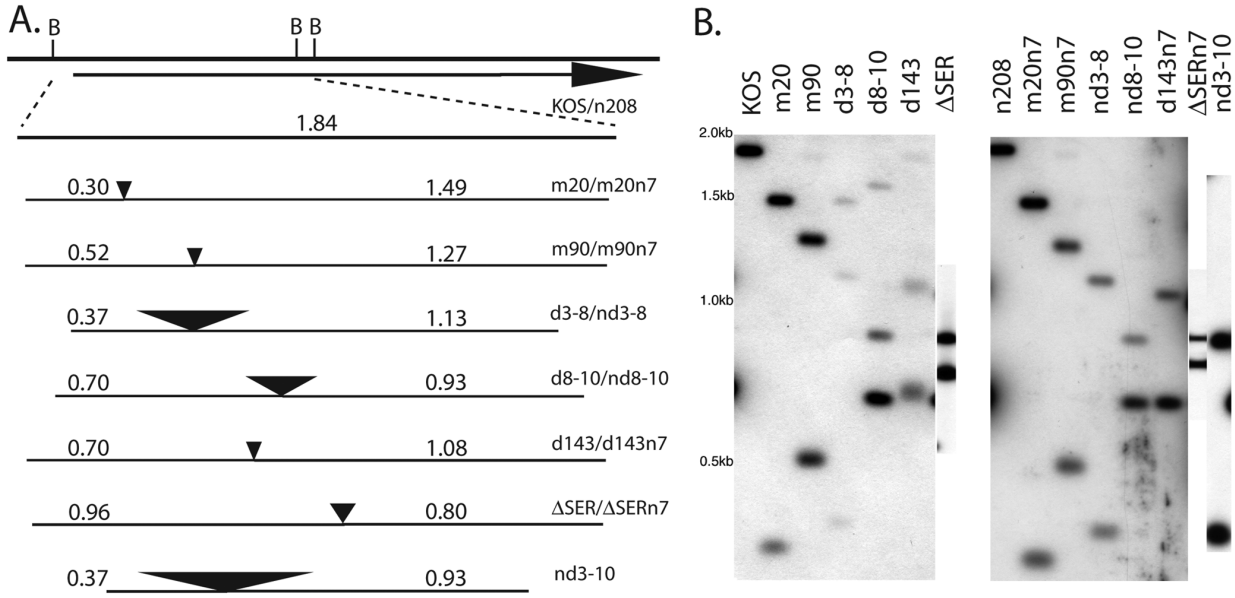


Figure 11. Structure of mutant virus genomes.
(A) Diagram of deletions relative to the ICP4 gene. The location (PstI site) and size of the deletions are given by the vertical arrowhead. The sizes (kb) of the predicted BamHI/PstI fragments for each mutant are also given. (B) Southern blot of BamHI/PstI digested viral DNA. The 1.84 kb BamHI "Y" fragment (part A) was used as probe.

Once it was established that each viral mutant had the correct genotype, we wanted to ascertain; (i) that each mutant virus was expressing an ICP4 protein of the correct size, (ii) whether each mutant molecule localized to the nucleus of infected cells, and (iii) if mutant molecules retained their ability to bind to DNA. To investigate whether mutant ICP4 molecules were expressed and whether the molecules were the expected molecular weight, SDS polyacrylamide gel electrophoresis (SDS-PAGE) followed by Western blot analyses were performed. The analysis was performed using the ICP4 N-terminal N15 polyclonal rabbit serum as a probe. Fig. 12A shows the results of the Western blot analyses. The mutants, nd3-8 and nd3-10, appear as though they are not expressed within the context of viral infection. However, as these mutations consist of large deletions within the N-terminus and the antibody N15

recognizes regions within the N-terminus, we hypothesized that the N15 antibody primarily recognized the regions deleted in d3-8 and nd3-8. To address this concern in part, Western blot analyses were also performed using the ICP4 C-terminal monoclonal antibody, 58S. As also displayed in Fig 4A, the d3-8 mutant is expressed to a similar level as wild type ICP4. While the antibodies at our disposal cannot detect nd3-8 and nd3-10, we presume that these truncated proteins are being produced. Additionally, all deletion mutants produced ICP4 molecules of the anticipated molecular weight relative to wild type ICP4 (Fig. 12A)

Immunofluorescence assays were performed to determine whether the mutant ICP4 molecules localized to the nucleus of the infected cells. Vero cells grown on glass coverslips were infected at an MOI of 10 for two hours before being fixed and stained with the ICP4 N-terminal antibody N15 and an AlexaFluor conjugated secondary antibody. Staining was imaged on a confocal microscope at 102x magnification. Figure 12B shows the images collected from the immunofluorescence assay. Images show ICP4 staining within the nucleus of infected cells indicating that all mutant ICP4 molecules localize to the nucleus during infection (Figure 12B). Again, the d3-8, nd3-8, and nd3-10 are not strongly recognized by this antibody. Some faint nuclear staining can be visualized with the d3-8 and nd3-8 mutants, however, staining is weaker than that of wild type virus.

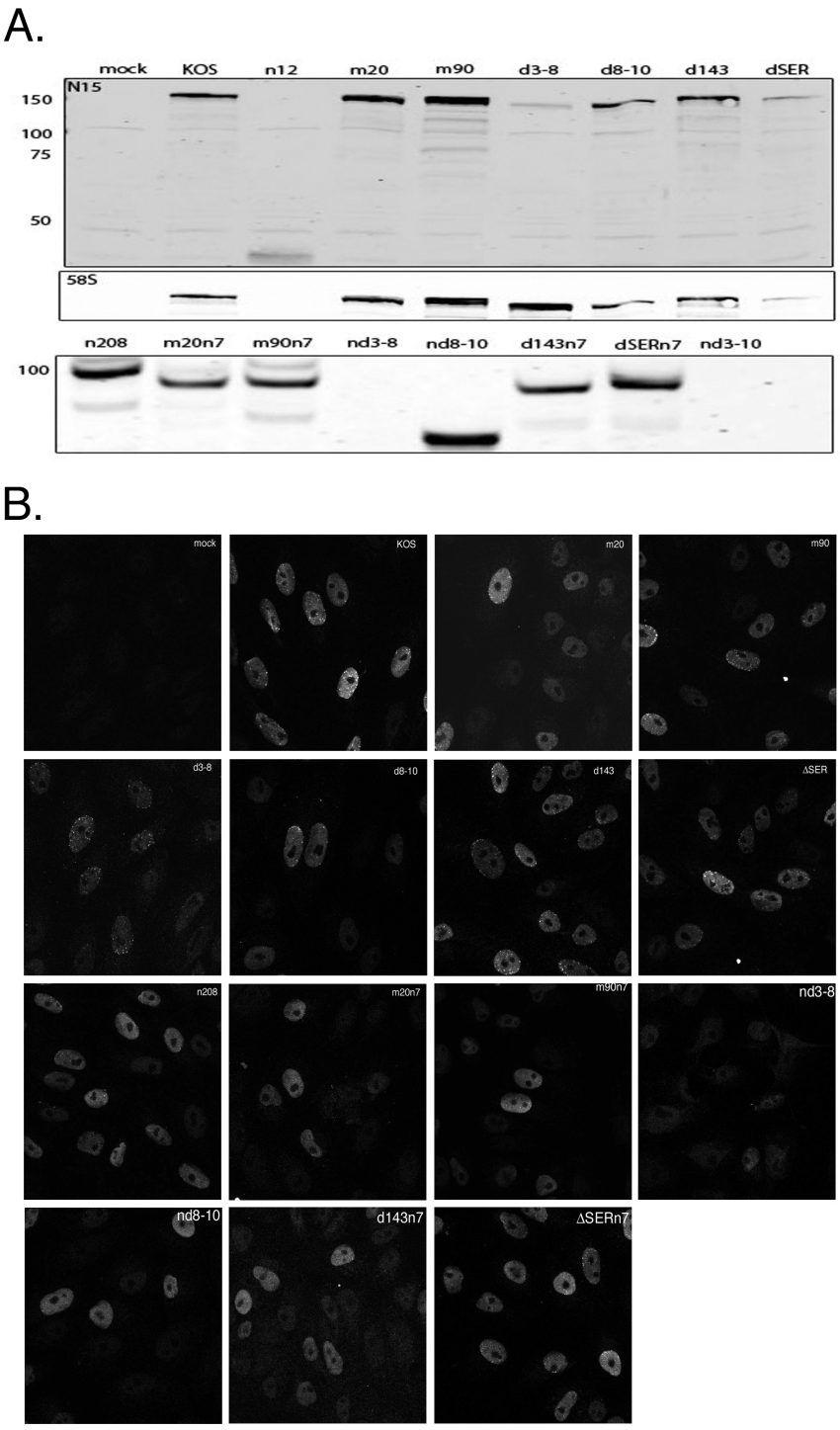


Figure 12. ICP4 mutant viruses produce proteins of the expected molecular weight and sub-cellular localization.

(A) Whole cell extracts were collected from wild type and mutant infected Vero cells 6hpi and proteins were separated on a 9% polyacrylamide gel. ICP4 proteins were detected with either the

ICP4 N-terminal N15 antibody (panels 1 and 3) or the ICP4 C-terminal antibody 58S (panel 2). (B) Immunofluorescence of ICP4 in virus infected cells.

To determine whether the mutant ICP4 molecules were capable of binding DNA, electromobility shift assays (EMSA) were performed. Protein extracts were collected from infected Vero cells and incubated with ^{32}P end-labeled DNA corresponding to the region -108 to +27 bp relative to the transcription initiation site of ICP4 in the presence of competitor DNA. Protein/DNA complexes were then separated on a 4% native polyacrylamide gel and the protein complexes bound to the radiolabeled probe were visualized by autoradiography. The results of the electromobility shift assay are in Figure 13. Bands were seen in the mock-infected lane indicating that some cellular proteins are capable of binding the ICP4 promoter region. An additional band is visible within the KOS infected extract band indicating that wild type ICP4 can bind the probe under these conditions. Additionally, all the ICP4 deletion mutants, with the exception of n12, that is missing the DNA binding domain, retained the ability to bind to the probe DNA as revealed by an additional band of the anticipated mobility present that is not present in the mock infected cell extracts (Fig. 13). Of importance, both nd3-8 and nd3-10, which were not detectable by western blot, also bound DNA, indicating that these mutant proteins are in fact produced in the context of viral infection. In addition, all of the mutants that possess the C-terminal 520 amino acids produced gel shifts of lower mobility and intensity relative to the main gel shift, while the viruses lacking the C-terminus only produced the single gel shift. This is due to the previously shown property of the C-terminus, which results in the multimerization of ICP4 on DNA (183).

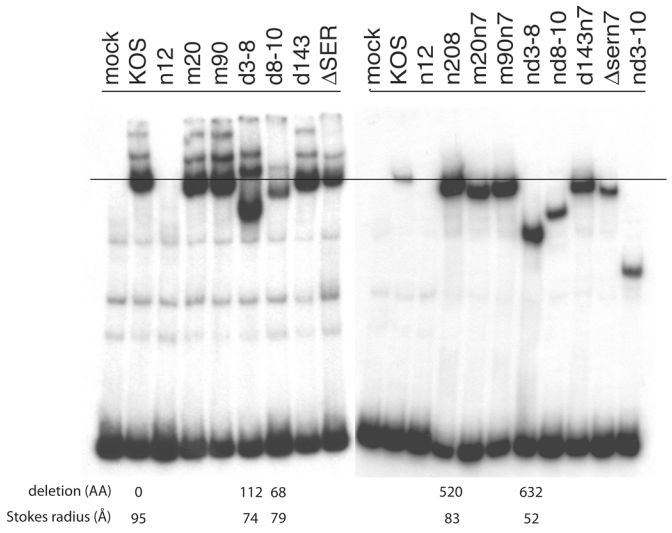


Figure 13. EMSA on ICP4 mutants.
Infected Vero cell extracts were prepared and incubated with ^{32}P end labeled DNA corresponding to the mRNA start site region of ICP4 as described. The protein/DNA complexes were separated by native PAGE and visualized by autoradiography. Each lane is labeled with the corresponding virus in which the protein extracts were collected from.

An additional property of ICP4 is revealed by the mobility of the shifts in Fig. 13. Deletions within the N-terminus significantly increased the mobility of the ICP4/DNA complex, whereas deletion of the entire C-terminal 520 amino acids (n208) had a relatively small effect by comparison. Size exclusion chromatography on superose 6 was performed on the full-length protein (KOS), n208, d3-8, d8-10 and nd3-8 ICP4 molecules to determine their Stokes radii. Relatively small deletions from the primary sequence at the amino terminus (d3-8 and d8-10) resulted in greater reductions in the Stokes radius than deletion of the 520 amino acids at the C-terminus. The reduction in Stokes radii roughly corresponded to the increase in the mobility of the protein/DNA complex. These data suggest that the amino terminus more greatly contributes to the elongated nature of the molecule.

The fact that all ICP4 mutant molecules produced the expected size protein, retained the ability to bind to DNA, and localized to the nucleus indicates that any defects in the regulation of viral gene expression are a direct effect of the deletions within the transactivation domain(s). To

determine how the deletions affect the virus life cycle, single step growth and viral gene expression analyses were performed on mutant virus infected cells. For single step growth analysis, Vero cells were infected at an MOI of 5 for 2, 4, 8, 12, 24, and 36 hours. Cells were harvested in their media, subjected to three freeze-thaw cycles, sonicated, and viral lysates were clarified via centrifugation. Viral yields were determined by plaque assay on the E5 complementing cell line and plotted versus time post infection (Fig. 14).

At 24 hpi wild type KOS infected cells yielded approximately 10^9 PFU, which is approximately 1000 PFU per cell. The DNA binding deficient virus n12, which contains only the first 251 aa of the ICP4 protein, did not produce a viral yield higher than the initial inoculum and was reduced by more than 4 orders of magnitude with respect to wild type. The viral growth profiles of m20, m90, d143, and ΔSER did not differ significantly from wild type. These mutants all contain small, approximately 10 aa, deletions within the N-terminus of ICP4. However, the mutants d3-8 and d8-10, which have deletions of 112 and 68 amino acids respectively, were reduced in viral yield by approximatly 1.5 orders of magnitude (Figure 14A). The yield of the C-terminal deletion mutant n208 was reduced by approximately 2.5 orders of magnitude, relative to wild-type virus. In the absence of the C-terminus the m20n7 and d143n7 mutants have growth profiles similar to n208 while the m90n7, Δsern7, nd3-8, nd8-10, and nd3-10 had growth profiles similar to the n12 virus (Figure 14B). Of these m90n7 and Δsern7 are perhaps the most interesting as in the presence of the C-terminus they have growth profiles similar to wild type, but in the absence of the C-terminus appear to have a relatively greater effect, underscoring the importance of these regions when the C-terminus is not present. Furthermore, the relative growth properties of the mutant viruses closely correspond to the

ability of the mutant plasmids from which they were derived to complement d120, strongly supporting that the defects of the viruses are due to the mutations in ICP4.

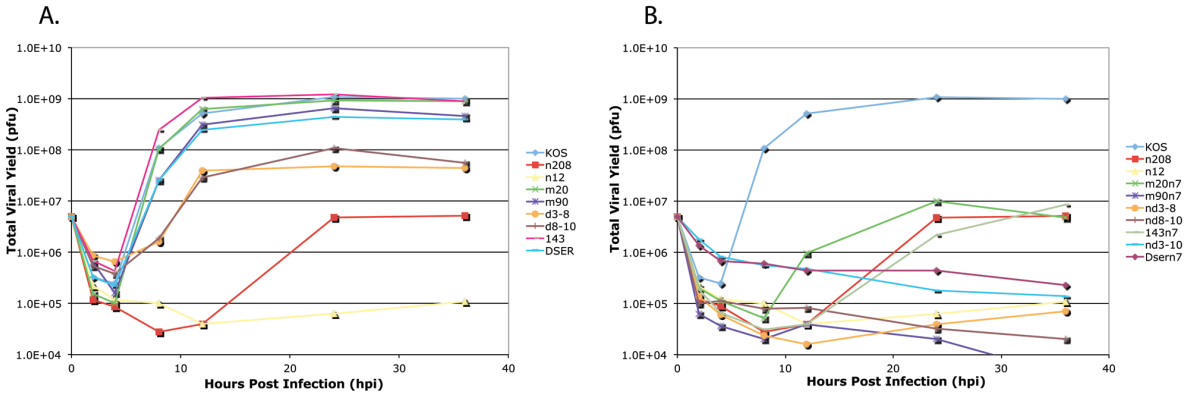


Figure 14. Single step growth curve analysis of mutant viruses.
(A) Growth curve analyses of ICP4 mutants that contain the C-terminus. (B) Growth curve analyses of ICP4 mutants that do not contain the C-terminal activation domain.

Given that the deletion mutants displayed varying defects in viral growth (Fig. 14), it may be anticipated that the spectrum of viral genes expressed as a function of the mutant ICP4 molecules will also vary, and differ considerably from that of wt virus. To investigate this, the kinetics of viral polypeptide synthesis, and the abundance of representative viral mRNAs was determined for the mutants and wt virus. To investigate viral protein synthesis, Vero cells were infected at a MOI of 10 PFU/cell with the indicated viruses, and the proteins were metabolically labeled with ^{35}S-methionine at 3.5-4, 7.5-8, and 11.5-12 hours post infection. Laemmli extracts of the infected cells were prepared, and subsequently separated on 9% SDS-polyacrylamide gels. Fig. 15 is a collage of SDS-gel profiles of mock-, wt virus-, n208-, m90n7-, nd3-10-, and n12-infected cells. The phenotypes of these viruses are representative of five general classes (I-V). The phenotypes of the remainder of the viruses are categorized into one of the five classes, and indicated below the representative member of the class.

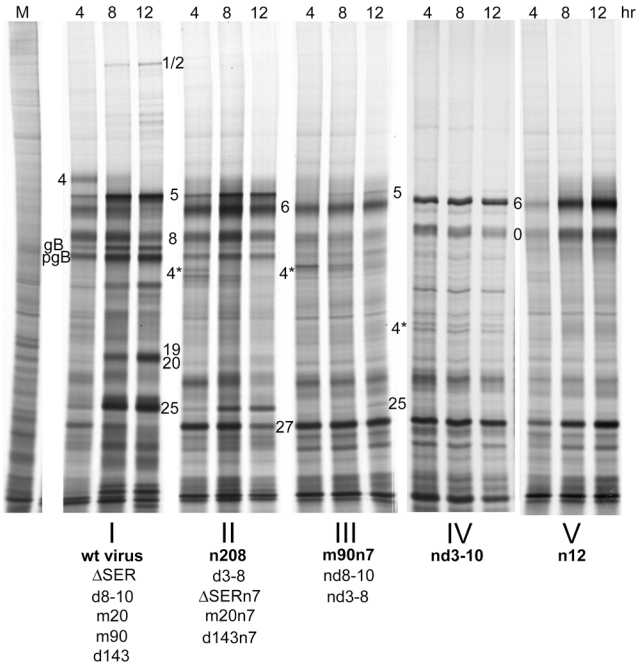


Figure 15. Protein expression profiles from ICP4 mutant viruses.
Vero cells infected with the indicated viruses were labeled for 30 min at 4, 8, and 12h post infection. SDS-PAGE was performed as described. Viruses with similar phenotypes are listed below the examples shown. Where possible infected cell polypeptide numbers (ICPs) numbers are given. Mock infected cells (M).

Class I is represented by wt virus (strain KOS). The classic progression from IE (ICPs4, 27) to true L (ICPs1/2, 19/20) is evident. The viruses, ΔSER, d8-10, m20, m90, and d143 share this phenotype, suggesting that these mutations alone do not greatly affect the ability of ICP4 to function. In contrast, class V, which is represented by the virus n12, was completely defective. The proteins ICPs0, 6, and 27 were highly overexpressed and later proteins were not easily detected, as previously shown (70). None of the mutants in this study precisely displayed the phenotype of n12.

Class II is represented by n208. n208 has nonsense mutations in all three reading frames such that the molecule is truncated at amino acid 774. Despite lacking the C-terminal 520 amino acids, readily detectable quantities of ICPs5, 8, gB, and 25 were detected, while the expression of the mutant ICP4 protein and ICP27 diminishes over time. Importantly true late proteins were not

easily detectable. This is consistent with previous findings (70). Like their counterparts in the context of the otherwise wild type background, the mutations in ΔSERn7, m20n7, and d143n7 had little effect on the n208 phenotype. Interestingly, the 112 amino acids in the N-terminus of the d3-8 ICP4 (Fig. 9) had similar effects on the activity of the ICP4 molecule as the lack of the C-terminal 520 in n208, as determined by this assay.

Class III is represented by m90n7. This 15 amino acid deletion (Fig. 9) had a profound effect on the activity of the n208 molecule, greatly reducing the abundance of ICPs5, 8, 25, and gB. The region deleted by this mutation represents the most conserved stretch of amino acids in the N-terminus of ICP4. Interestingly, this deletion had a relatively small effect in the presence of the C-terminal region of ICP4 (m90). nd3-8 had a similar phenotype to m90n7, consistent with the fact that the region deleted in nd3-8 contains the region deleted in m90n7. The deletion in nd8-10 consists of 68 amino acids just C-terminal to the d3-8 deletion residues, and has a similar effect on the activity of the n208 molecule. Thus two non-overlapping regions of the N-terminus have a similar effect on the activity of the ICP4 molecule.

Class IV is represented by the virus nd3-10. It is the only member of this class. The deletion in this virus combines the two deletions that independently resulted in a marked reduction in viral later gene expression in the n208 background. As a consequence gene expression beyond the IE class of virus genes was very difficult to detect and similar to class V (n12). The phenotypes nd3-10 and n12 subtly differ in that ICP6, 0 and 27 appear to be more abundantly expressed in n12-infected cells.

To more directly assess the activities of the mutant ICP4 molecules synthesized during infection, the abundance of mRNAs for genes representative of different kinetic classes was determined by Northern blot analyses (Fig. 16). The abundance of ICPs4 and 0 (IE), tk (E), and

gC and LAT (L) mRNAs were determined at 4 and 8 h post infection. The IE transcripts for ICP4 and ICP0 accumulate differently during infection. ICP4 accumulates early and then is reduced in abundance, whereas ICP0 continues to accumulate throughout infection (351). The decreased accumulation of ICP4 is due the ability of ICP4 to repress transcription as a function of an ICP4 binding site at the start site of ICP4 transcription. All of the mutants show less ICP4 mRNA at 8h relative to n12, suggesting that they all may retain the ability to repress transcription. Another observation that supports this is that none of the mutants that express LAT demonstrate accumulation at early times. It has been shown that the LAT promoter is converted to an early promoter by mutating a repressive ICP4 binding site at the start site of LAT transcription (271). Interestingly, in the absence of the C-terminus of ICP4, ICP0 is still highly expressed at 8hpi, but is not as highly expressed as wild type. This is true of all mutants lacking the C-terminus (Figure 16, panel 2), perhaps reflecting the defect in activation.

tk (early) mRNA was not detected at this level of exposure in n12-infected (Fig. 16). At 4h post infection, the abundance of tk mRNA in n12-infected cells is about 2% that of wt (strain KOS) virus (140). While tk RNA was expressed in all the backgrounds containing an intact C-terminus, its abundance was significantly reduced in m90- and d3-8-infected cells. The larger growth defect (Fig. 14) and more restrictive SDS-gel profile (Fig. 15) of d3-8 relative to m90, is probably because of defects in later gene expression. The abundance of the late RNAs (gC and LAT) in d3-8 infected cells were considerably less than in m90-infected cells (Fig. 16). In the absence of the C-terminus, the abundance of tk RNA in nd3-10 infected cells was as low as in n12-infected cells. tk mRNA was also significantly reduced in m90n7- and nd3-8-infected cells. n208 was defective in late gene expression as previously shown (70), as were all the viruses lacking the C-terminal 520 amino acids. Some of the other deletions had a more modest effect

on tk and gC mRNA abundance. The deletion of amino acids 142 to 210 in d8-10 resulted in a modest reduction in tk, gC, and LAT accumulation. Lastly, the deletion in d143 affected the kinetics IE (ICP4 and ICP0) and tk expression. Therefore, while the conserved block of amino acids deleted in m90 (81-96) (Fig. 9) had the greatest effect on the accumulation of the analyzed transcripts that ICP4 activates, determinants flanking this block also affect activation as inferred by the greater defect in d3-8 relative to m90. The region deleted in d8-10 further contributes to activation, as inferred by the highly defective phenotype of nd8-10.

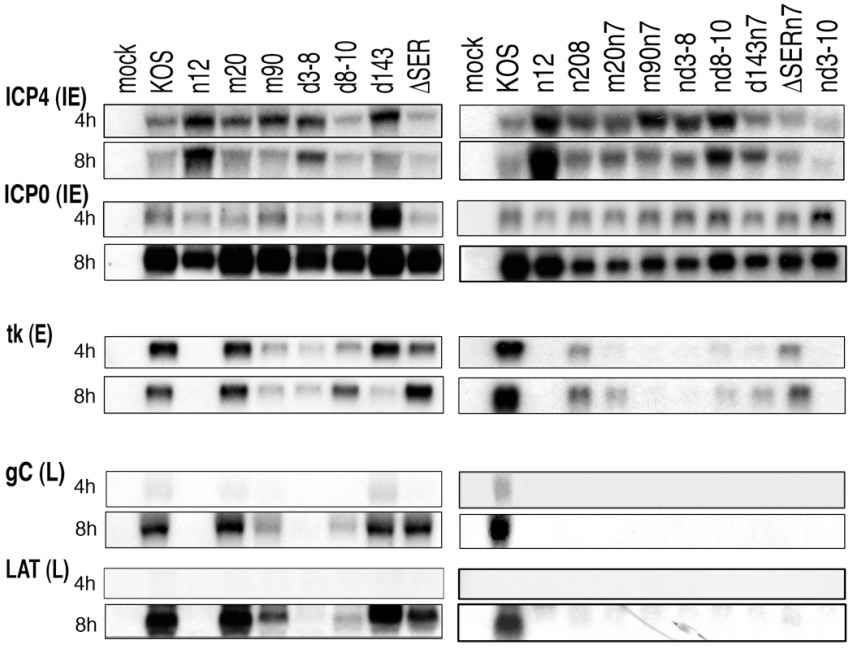


Figure 16. mRNA expression from ICP4 mutant viruses.
RNA was isolated from cells infected at an MOI of 10 PFU/cell for 4 and 8 h. Northern blot analysis was performed as described in the materials and methods, probing for ICP0, ICP4, tk, gC, and LAT.

2.5 DISCUSSION

This study addresses the extent to which the conserved regions of the N-terminus that contribute to the regulatory functions of the molecule, and whether the N- and the C-termini of ICP4 have any redundant or overlapping functions. To address these questions, directed deletions were constructed within the N-terminus of ICP4, both in the presence and absence of the C-terminal regulatory region and the resulting mutant viruses were compared with respect to viral gene expression. We have previously shown that the region N-terminal to the DNA binding domain of ICP4 is important for activation and repression (41, 119, 295). In this study we found that multiple regions within the N-terminus contribute to activation. To put these findings into perspective it is helpful to consider some general properties of the ICP4 molecule and the N-terminus in particular.

ICP4 exists in cells as an obligate dimer possessing hydrodynamic properties suggestive of an elongated protein with a Stokes radius of about 95-100 Å (223, 296). The N-terminus contributes greatly to the elongated character of ICP4 relative to the C-terminus (Fig. 13). Deletions in the 50 to 100 amino acid range within the N-terminus more substantially reduced the Stokes radius of ICP4 compared to the deletion of the entire 500 amino acids of the C-terminus. The N-terminus also has a greater effect on the magnitude of the mobility shift of an ICP4/protein DNA complex (Fig. 13). This suggests that an ICP4 molecule bound at a fixed ICP4 binding site may be able to act at a considerable distance engaging in interactions between cellular molecules and domains within the N-terminus.

Another property to consider is the predicted degree of order vs disorder in the ICP4 molecule. Disordered regions are generally of low complexity and are flexible. The probability of disorder can be predicted (345) with the use of an online server,

http://bioinf.cs.ucl.ac.uk/disopred (344). Figure 17A shows this prediction for the entire ICP4 (strain KOS) protein. There are two regions that have a fairly low probability of disorder. The first is the domain between amino acids 300 and 600. This constitutes the DNA binding/dimerization domain. The second is the C-terminus of the molecule roughly from amino acids 900 to 1294. This is largely the region deleted from the n208 molecule. It is also the region that when deleted had a relatively small effect on the Stokes radius. Together these observations suggest that this is a relatively ordered, more globular part of the ICP4 molecule. In between the 2 regions of relative predicted order is a region from amino acids 700-850 that is predicted to be highly disordered. This region is poorly conserved among the ICP4 homologs in other α-herpesviruses.

The amino terminus is a heterogeneous mixture of ordered and disordered regions, with the conserved regions not necessarily corresponding to the degree of predicted disorder (Fig.17). A conserved disordered region may fold upon binding to a functional target, such as a cellular transcription factor. Other disordered regions may serve as flexible linker regions between interacting regions (83). One such region may be that corresponding to amino acids 82-97, which are deleted in m90 (Fig. 17C). In the absence of the C-terminus, this region had a relatively large effect on the activity of the molecule. It is also the most conserved region within the region within the N-terminus surveyed by this analysis (Fig. 17). It is possible that this region is important with making contacts with key transcription factors such as TFIID and mediator, which are known to form a complex with ICP4 in infected cells (198). However, the participation of any region within the N-terminus in contacts with other molecules cannot be ruled out. It is likely that there are multiple regions within the N-terminus that participate in protein-protein interactions that are connected by flexible linker regions, enabling ICP4 to make

multiple contacts with the cellular transcription machinery, possibly at a distance from where it is bound to DNA.

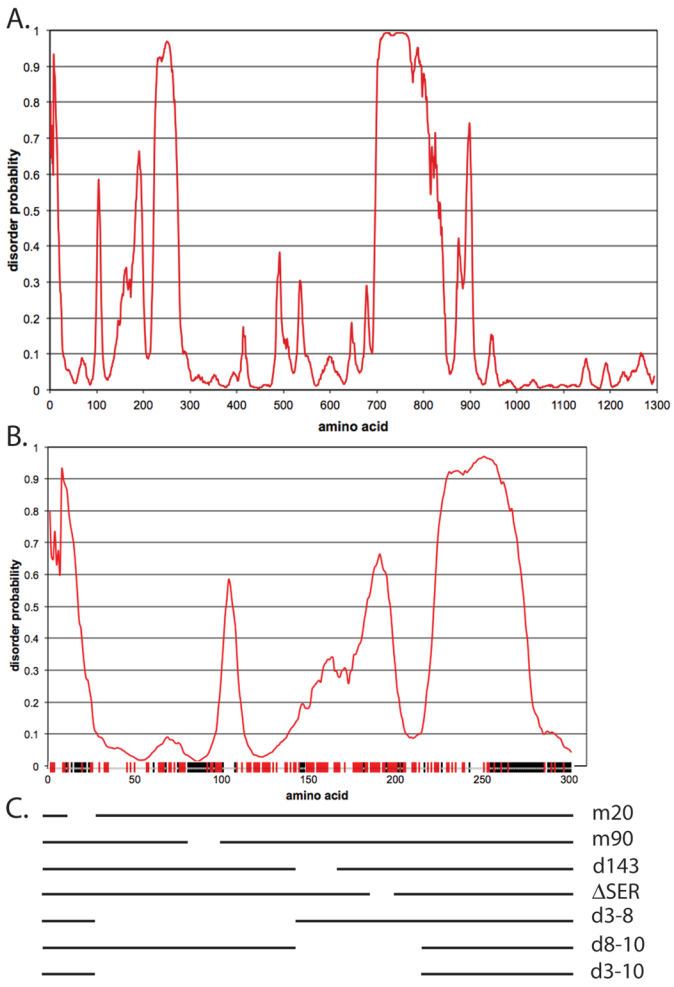


Figure 17. Disorder prediction for the ICP4 protein.
(A) The probability of disorder for ICP4, (strain KOS) was determined as described in the text. (B) Expansion of part A for the amino terminus of ICP4. The degree of conservation between HSV-1, HSV-2 and B virus from Fig. 9 is also shown along the axis depicting the residue number. (C) The relative locations of the deletions examined in these studies.

The presence of the C-terminus almost completely masked the deleterious effects of the m90 deletion. Therefore, the C-terminus independently substitutes for the activity of the deleted amino acids, or it augments the activity of the N-terminus that remains when these amino acids

are deleted. The n208 molecule interacts with TBP and TFIID on DNA (117, 303), and the contributing protein-protein interactions are specified by the amino acids spanning 30-274 (117, 119). However, the C-terminus of ICP4 specifies an interaction with TAF1 enabling ICP4 to interact with TFIID in solution (41). Therefore, there are regions within both the carboxyl- and amino-terminal regions that specify interactions with TFIID that could potentially synergize to recruit TFIID. The regions of ICP4 involved in interactions with mediator are unknown at present. However, since mediator is ubiquitously involved in polII transcription, as is TFIID, regions interacting with these two central complexes will be operationally redundant to some extent.

The region deleted in d8-10 (142-210) had a considerable effect on activation, particularly in the absence of the C-terminus. Part of this is due to contribution of the region deleted in d143. Because d8-10 still retains sequences deleted in m90, it likely functions in gene activation by a distinct mechanism. At present it is not clear whether this contribution is as a spacer or due to interactions with different transcription factors, or both.

It is likely that the entire region comprising the N-terminal 200 amino acids is involved in interactions with multiple transcription factors. The hydrodynamic properties of mutants in this region suggest that it extends out from the DNA binding domain, and is potentially flexible, allowing it to contact multiple transcription factors and act at a relative distance from where ICP4 is bound to DNA. This region, along with the DNA binding domain functions as an activator, resulting in the activities characteristic of n208. It is important to note that while n208 is relatively defective for viral growth compared to wt virus, some viral growth does occur (Fig. 14). Perhaps the more globular and conserved C-terminal region (amino acids 774-1294) "augments" the activities of the N-terminus either by specifying redundant interactions with the

N-terminus, or interactions affecting the location of ICP4. One example of the latter type of interaction is that the C-terminal 500 amino acids allow ICP4 to multimerize on DNA, increasing its affinity for DNA (183). This would enhance the ability of the interactions specified by the N-terminus to contribute to activation. Further studies are underway to test this model.

2.6 ACKNOWLEDGEMENTS

I would like to thank J.T. Lester for pioneering these studies and F.L.Sivrich for the construction of several viral mutants.

3.0 A FUNCTIONAL RELATIONSHIP BETWEEN HSV-1 TRANACTIVATION DOMAINS FACILITATES INTERACTIONS WITH TFIID AND MEDIATOR

3.1 ABSTRACT

The HSV IE protein, ICP4, is crucial in regulating viral gene expression; in its absence E and L gene expression do not occur. Multiple studies have defined regions of ICP4 that are important for viral gene expression. The C-terminus of ICP4 plays a critical role in activating L gene expression. The previous chapter suggested that the N-terminus is important for Early gene expression, although the specific functions were not defined. For the purposes of this study, a stable cell line expressing a mutant ICP4 deleted for the N-terminal transactivation domain, comprised of amino acids 30 through 210 (d3-10), was created and termed D14. Infection of d3-10 expressing cells with an ICP4 null virus yielded no viral progeny and showed a severe deficiency for Early mRNA and protein expression as assessed by qRT-PCR and SDS-PAGE analyses, respectively. The defects observed in transactivation of E genes were a direct result of the loss of functions executed by the N-terminus, as the DNA binding and localization functions remained intact. The absence of early gene expression coincided with a defect in the recruitment of RNA polymerase II to Early promoters. Interestingly, immunofluorescence data indicated that the mediator complex was not recruited to sites of ICP4 deposition, while TFIID was. Perhaps most interestingly, while TFIID localized to sites of ICP4 deposition, it did not form a complex

with d3-10 in affinity purification experiments. Additionally, when d3-10 was combined with the ICP4 mutant n208, which is missing the C-terminal transactivation domain, the defects observed were complemented and viral progeny were produced. This was the result of the formation of functional heterodimers, indicating that one functional N- and C-terminal transactivation domain within a dimer is sufficient for ICP4 mediated gene expression.

3.2 INTRODUCTION

ICP4 of HSV-1 is a 1294 amino acid protein consisting of two transactivation domains separated by a DNA binding domain and a nuclear localization sequence (253). It is well established that the DNA binding activity of ICP4 is crucial for activation and repression (253, 254, 295). The N-terminal and C-terminal transactivation domains mediate transcription, in part, by interacting with several host transcription factors including TBP, TAF1, and the Mediator complex (41, 117, 198, 303). It is presumed that through these interactions, RNA Polymerase II is preferentially recruited and stabilized on viral promoters, thus driving the initiation of viral transcription (284).

Genetic and biochemical studies of ICP4 indicate that the protein has a modular structure and that individual domains contribute to the activities of the protein differently. Expression of an ICP4 mutant lacking the C-terminal activation domain, n208, displayed reduced levels of E gene transcription and biochemically undetectable levels of L gene transcription (70, 295). Chromatin immunoprecipitation assays showed that RNA PolII and TBP E promoter occupancy was decreased and L promoter occupancy was undetectable in n208 infected cells, probably accounting for the reduced levels of E and L gene expression (284). Interactions with TAF1 of TFIID have been mapped to the C-terminus of ICP4, suggesting that the transcriptional defects

observed in n208 infected cells were the result of the molecule being less able to stabilize TFIID onto promoters (41). The additional removal of the N-terminal transactivation domain yielded a viral mutant (X25) that was not capable of IE gene repression or E or L gene transcription, but retained the ability to bind to DNA, suggesting that the C-terminus of ICP4 augmented activities specified by the N-terminus (296). The activities of the C-terminal transactivation domain in the absence of the N-terminal transactivation domain have not been determined.

The N-terminus of ICP4 is structured such that it is largely disordered with small conserved structurally ordered regions throughout (340). Deletion of these regions demonstrated that they contribute to the transcriptional activities of ICP4. The additional removal of the C-terminal transactivation caused detrimental defects to viral growth, demonstrating that the N-terminus and C-terminus of ICP4 contain redundant functions (340). Additionally, these studies suggested that regions in the N-terminus may have combined activities, possibly due the contributions of these regions in forming contacts with multiple transcription complexes.

To address whether the N-terminal transactivation domain contained regions specifying for combined activities and to determine the sufficiency of the C-terminal transactivation domain, a deletion mutant (d3-10) was constructed that deleted amino acids 30-210 from the amino terminus of ICP4. This region contains the majority of the proposed functional region of the N-terminus of ICP4, without interfering with regions defining DNA binding activities (254, 295). Biochemical analyses indicated that d3-10 is defective in activating E gene transcription, probably the result of insufficient recruitment of RNA polymerase II to early promoters as a consequence of the inability of the mutant molecule to form complexes with TFIID and possibly Mediator. Additionally, the observed defects of d3-10 could be complemented by the addition of a molecule specifying for the N-terminal 774 amino acids. These data suggest that the C-

terminal activation domain is not sufficient to support viral transcription and that the N-terminal and C-terminal transactivation domains contain combined activities.

3.3 MATERIALS AND METHODS

Cells and Viruses. Vero cells (African Green Monkey Kidney cells) were cultured in Dulbecco's Modified Eagle Medium with 5% FBS and maintained as suggested by ATCC. E5 cells are stably transformed with a wtICP4 encoding plasmid, pk1-2 (69), and psV2neo and have been describe previously (69). D14 cells were constructed by transfecting Vero cells with 6ug of a plasmid containing the d3-10 mutant and 2 ug psV2neo (see below). I3 cells were constructed by transfecting Vero cells with 6 ug of a plasmid encoding TAP-d3-10 and 2 ug psV2neo (see below).

The viruses KOS, d120 (67), n208 (70), and n12 (70) have been described previously and with the exception of KOS, were propagated on E5 cells. Briefly, d120 has the majority of the ICP4 coding sequence deleted, n12 expresses only the first 251 aa, n208 expresses only the first 774 aa, and KOS is the wild type strain used in these assays.

Production of stably expressing d3-10 and TAP-d3-10 cell lines. 2 x 10^6 Vero cells were transfected with 2 μg psw2neo and 6 μg d3-10 or TAP-d3-10 plasmid DNA using 15 μL lipofectamine 2000 according to the manufacturers suggestion (Invitrogen). The d3-10 plasmid contains the ICP4 d3-10 deletion mutant driven by its own promoter (295). The TAP-d3-10 plasmid contains a streptavidin binding peptide and a calmodulin binding peptide (Stratagene) N-terminal to the d3-10 ICP4 under control of its endogenous promoter. Two days post transfection the cells were trypsinized, counted, and approximately 10,000 cells were plated on

100 mm dishes in DMEM with 10% FBS. The following day the medium was replaced with DMEM containing 10% FBS and 800 μg/mL G418 (Invitrogen), allowing for selection of neomycin resistant cells. Medium was replaced every 3-4 days for 17 days until sufficient colonies of cells were visible. Forty colonies were isolated from 9 plates of cells, expanded, and characterized for d3-10 protein expression. Eighteen colonies were isolated from 10 plates of cells, expanded, and characterized for TAP-d3-10 expression. Cultures from isolated colonies were maintained in DMEM with 5% FBS and 500 μg/mL. Nine of the 40 and 5 of the 18 clones were positive for expression of the d3-10 or TAP-d3-10 protein, respectively, as determined by infection with the ICP4 null virus d120 followed by Western blot analysis. An isolate expressing comparable level of the d3-10 protein to wtICP4 from KOS infected cells was utilized for the assays described herein. The isolate expressing the most TAP-d3-10 (I3) was utilized for the affinity purification assays.

SDS PAGE and Western Blot Analyses. SDS-PAGE assays and Western blot analyses were performed as previously described (340). Briefly, 10% polyacrylamide Tris-HCl gels (Biorad) were loaded with samples diluted in Laemmli buffer (0.05 M Tris-HCl, 2% SDS, 5% β-mercaptoethanol, 0.01% bromophenol blue, 5% sucrose). Electrophoresis was carried out using Tris-Glycine Electrode buffer pH 8.5 (25 mM Tris-HCl; 192 mM Glycine; 1% SDS) at 140 V for 1.25 hours. Samples subjected to Western blot analyses were transferred to a nitrocellulose membrane in transfer buffer (25 mM Tris-HCl; 192 mM Glycine; 20% methanol) at 250 mA for 1.5 hours. Immunoblots were blocked in a 5% milk TBS-T (50 mM TrisHCl, 150 mM NaCl, 0.05% Tween20; pH 7.5) overnight. Immunoblots were incubated with primary antibodies for ICP4 (58S; 1:500) for 1hour at ambient temperature. Immunoblots were washed 4 times for a total of 30minutes in 1% milk TBS-T, then incubated with IRDye conjugated secondary

antibodies (LICOR) in TBS-T (1:10,000) for 1hour, rinsed 4 times for a total of 30 minutes in TBS-T, twice for 5 minutes in TBS (50 mM TrisHCl, 150 mM NaCl pH 7.5), and finally visualized using the LI-COR Odyssey Infrared Imager.

Metabolic Labeling of Polypeptides. Metabolic Labeling experiments were performed as previously described (54, 340) with the following modifications. Five hundred thousand Vero, E5, or D14 cells were infected at an MOI of 1 with KOS or an MOI of 10 with d120 for 1hr at room temperature. The infections were allowed to progress at 37° for 3.5, 7.5, 11.5, or 23.5hrs at which time the cysteine and methionine-free medium was removed and cells were incubated with 22 μCi ^{35}S-Methionine (>1000 Ci/mMol, MP Biomedical) in TBS for 30minutes at 37°. Labeled proteins were separated through a stacking gel (4% acrylamide, 0.2% DATD, 0.1% SDS, 119 mM Tris, pH7.0) and then through a separating gel (9% acrylamide, 0.2% DATD, 375 mM Tris, 0.1% SDS, pH 8.8) at 100V. The gel was fixed in fixing solution (water: methanol: acetic acid at a 6:3:1 ratio) and subjected to autoradiography.

Immunofluorescence Assays. Two hundred thousand Vero, E5, or D14 cells were grown on glass coverslips in 12-well trays. Cells were infected at the indicated MOIs (1 for PML experiments, 10 for all other experiments) for 1hr at room temperature. The inoculum was removed and fresh media was added. The infections were carried out at 37° for the indicated period of time. Cells were washed in PBS and then fixed in 4% PFA for 10 minutes. If the cells were detergent extracted, they were rinsed in PBS containing a 1X protease inhibitor cocktail (Roche), and extracted in extraction buffer (10 mM Hepes pH 7.4, 300 mM sucrose, 100 mM NaCl, 3 mM $MgCl_2$, 0.5% Triton-X-100) on ice for 2 minutes before fixation with PFA. Fixed cells were permeablized in PBS containing 1% BSA and 0.4% Triton-X-100 for 30 minutes. Cells were blocked in PBS containing 1% BSA for an additional 30 minutes. Cells were

exposed to primary antibodies for 1 hour [ICP4: N15 1:500, 58S 1:500; Trap220: (Santa Cruz, sc-8998) 1:250; TBP: (Santa Cruz, sc-273) 1:150], washed in PBS+ 1% BSA for 30 minutes, and washed in PBS 3 times for a total of 30 minutes. Cells were incubated with Alexa-Fluor conjugated secondary antibodies at a 1:500 dilution for 45 minutes before being washed 6 times for a total of 60 minutes in PBS and mounted to glass slides using Immu-Mount (Thermo-Fisher). Fluorescence was visualized using the Olympus Fluoview 1000.

Complementation Assays. 5 x 10^5 Vero, E5, or D14 cells were infected at an MOI of 5 for 1 hour at room temperature. The inoculum was removed, the cells were washed three times in cold TBS, and fresh media was added. At 24 hpi, infected cells were harvested in their own media, freeze-thawed three times, sonicated for 45 sec, and the viral lysates were clarified. Total viral yields were determined by plaque assay on E5 cells.

Affinity Purification. 5 x 10^5 Vero, E5, D14, or I3 cells were infected at an MOI of 10 with KOS, n208, TAP-ICP4, TAP-n208, or d120 for 1hr at 37°. The inoculum was removed, fresh media was added, and infected cells were placed at 37° for 5 hours longer. Cells were washed once with TBS + 0.1 mM TLCK, and scraped into 10mL TBS + 0.1 mM TLCK. Nuclear extracts were collected as previously described by Dignam (76). Briefly, the cells were pelleted at 3 K for 5 minutes, and resuspended in 5 mL Hypotonic buffer (10 mM Hepes pH7.9, 1.5 mM $MgCl_2$, 10 mM KCl). The cells were allowed to swell on ice for 10 minutes before being dounced. Trypan blue staining assessed the douncing efficiency. Nuclei were pelleted for 10 min at 3 K. Nuclei were resuspended in ½ the pellet volume low salt buffer (20 mM Hepes pH 7.9, 20% glycerol, 1.5 mM $MgCl_2$, 20 mM KCl) and ⅓ the total volume of high salt buffer (20 mM Hepes pH 7.9, 20% glycerol, 1.5 mM $MgCl_2$, 1.4 M KCl) was added dropwise. Extracts were rotated end over end for 30minutes at 4° and spun at 12.5 K for 30 minutes. The nuclear

extracts diluted with 2X low salt streptavidin buffer (20 mM Hepes pH 7.9, 20 mM KCl, 4 mM EDTA, 0.2% NP40) and were then combined with 100 µL streptavidin beads (Pierce) and rotated overnight at 4°. The beads were washed in Streptavidin Binding Buffer (SBB) (20 mM Hepes pH 7.9, 200 mM KCl, 2 mM EDTA, 0.1%NP40) 6 times for 10 minutes. 500 µL Laemelli buffer was added to the beads. These samples were subjected to SDS-PAGE and Western blot analysis.

RNA Extraction and Quantification. RNA extraction and quantitation experiments were carried out as previously described with modifications (101). Vero, E5, or D14 cells were infected at an MOI of 10 with the indicated viruses. RNA was extracted from the cells using the RNAqueous-4PCR Kit from Ambion and the included protocol.

One microgram of RNA was reverse transcribed to cDNA using the Retroscript Kit from Ambion following the included protocol. Briefly, 1 µg total RNA was combined with 2 µL Oligo(dT) in a final volume of 12 µL and heated to 85° for 3 minutes. The RNA-oligo(dT) mixture was combined with 2 µL 10X RT Buffer, 4 µL dNTP mix, 1 µL RNAse Inhibitor, and 1 µL MMLV-RT. The RNA was reverse transcribed at 42° for 1 hr. The MMLV-RT was inactivated by heating the reaction to 92° for 10 minutes.

Chromatin Immunoprecipitation. Chromatin Immunoprecipitation assays were carried out as previously described (99-101, 198, 284). Briefly, 5 x10^6 Vero, E5, or D14 cells were plated onto 100 mM dishes. Cells were infected with the indicated viruses at an MOI of 10 for one hour at room temperature. Following adsorption fresh DMEM + 5% FBS was added to each plate and the plates were incubated at 37° for 3 hours 15 min at which time the medium was removed and crosslinking reagent (1% formaldehyde, 0.15% sodium bicarbonate, in DMEM + 5% FBS) was added. The samples were crosslinked for 10 minutes at 37°. Formaldehyde was quenched by the addition of glycine to the medium to a final concentration of 125 mM. Cells were washed

twice in TBS and scraped into TBS. The cells were briefly pelleted, resuspended in 500 μL SDS-Lysis Buffer, (1% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.1) and incubated on ice for 30 minutes. The cells were sonicated 6 times for 10 sec each time at a 30% output using a probe sonicator. The cell debris was pelleted at 12.5 K for 15 minutes and the sonicated chromatin was diluted 11 times in ChIP dilution buffer (0.01% SDS, 1.1% TritonX100, 1.2 mM EDTA, 20 mM Tris-HCl pH 8.1, 167 mM NaCl). The samples were pre-cleared twice for two hours using 120 uL of protein A agarose + ssDNA beads (Milipore) each time. The pre-cleared samples were divided into five samples and the immunoprecipitations were performed overnight using 30 μL Protein A agarose + ssDNA beads and either 5 μL of 58S antibody for ICP4 and 2 μL 8WG16 antibody for RNA Polymerase II. Beads containing the precipitated protein/DNA were washed twice for 4minutes each in Low Salt Buffer (0.1% SDS, 1% Triton X 100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 150 mM NaCl), twice for 4minutes each in High Salt Buffer (0.1% SDS, 1% Triton-X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 500 mM NaCl), once for 4 minutes in LiCl Buffer (0.25 mM LiCl, 1% NP40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.1), and three times for 4 minutes each in TE Buffer (10mM Tris-HCl pH=8.0, 1mM EDTA). The samples were eluted from the beads in 250 μL of Elution Buffer (1% SDS, 0.1 M sodium bicarbonate) twice for 15 minutes each. Twenty microliters of 5 M NaCl was added to each 500 uL sample and samples were incubated at 65° overnight to reverse the crosslinks. Thirty-one microliters of a Proteinase K Mix (1:2:1 0.5 M EDTA: Tris-HCl pH 6.5: 20 mg/mL ProteinaseK) was added to each tube and samples were incubated at 55° for 2 hours. DNA was purified using phenol chloroform extractions. The purified DNA was precipitated with ethanol, resuspended in 60 μL DNAse/RNAse free water and analyzed using qPCR with primers for the tk promoter were 5'CAGCTGCTTCATCCCCGTGG and 5'AGATCTGCGGCACGCTGTTG.

qPCR. qPCR reactions were set up such that they contained 7.5 μL SYBER Green 2X PCR mix, a final concentration of 0.3 μM primers (see charts below), and 3 μL cDNA (from RNA experiments) or DNA (from ChIP experiments) in a total volume of 15 μL. The reactions were carried out in a StepOnePlus Real Time PCR machine from Applied Biosystems under the following conditions: 95° for 10 minutes followed by 40 cycles of 95° for 15 seconds and 60° for 1 minute. A melt curve was also included with the following conditions: 95° for 15 seconds, 60° for 1 minute followed by +0.3° to 95° for 15 minutes. For cDNA quantification the primers used were 5'ACCCGCTTAACAGCGTCAACA3' and 5'CCAAAGAGGTGCGGGAGTTT3' for tk and 5'ACTTAATCAGGTTGTTGCCG3' and 5'GAAGTTGTGGACTGGGAAGG3' for ICP4.

3.4 RESULTS

The HSV-1 protein ICP4 temporally regulates viral gene expression by promoting transcription of E and L genes while concomitantly downregulating transcription of IE genes. These functions can be attributed to the abilities of ICP4 to bind DNA through its DNA binding domain and to interact with cellular transcription factors through two transactivation domains. It has been established that each of the transactivation domains is required to different extents for activation and repression. According to earlier findings, regions within the N-terminus of ICP4 significantly contribute to E gene transcription (11, 69, 340). Additionally, non-overlapping regions of the N-terminus have detrimental, but different, effects on early gene transcription; deletion of amino acids 30-142 resulted in lower levels of E transcripts while deletion of amino acids 142-210 resulted in the delayed accumulation of E transcripts (340).

Based on these observations, the extent to which the N-terminal transactivation domain of ICP4 contributed to the activation of E gene transcription was investigated.

A mutant, d3-10, containing a deletion of amino acids 30-210 was constructed to investigate the contributions of the N-terminus of ICP4 to viral transcription. This region comprises the majority of the amino terminal transactivation domain without interfering with the DNA binding domain (9, 11) (Figure 18 A). Because constructing the deletion within the viral genome was not possible, probably due to the transdominant nature of the molecule (data not shown), a cell line stably transfected with a plasmid encoding d3-10 under control of the wt ICP4 promoter was constructed. Several clonal isolates were tested for expression of the d3-10 protein via Western blot analysis (data not shown) and the D14 clonal isolate was chosen for use in these studies because it expressed comparable levels of ICP4 protein as KOS infected Vero cells (Figure 18B). Results from Western blot analysis of samples collected from mock, KOS, or and ICP4 null virus, d120 (67), infected Vero, E5, or D14 cells, are shown in Figure 18B. The presence of immunoreactive bands at 175 and 125 kDa of similar intensities established that E5 cells and D14 cells, expressed similar levels of wild type or mutant ICP4, respectively, as KOS infected Vero cells.

To determine if deletion of amino acids 30-210 interfered with the ability of the molecule to localize to and bind viral genomes, immunofluorescence and chromatin immunoprecipitation assays were performed. It has been previously reported that viral genomes localize to PML bodies in infected cells and that ICP4 localizes juxtaposed to viral genome containing PML bodies (92, 142). Accordingly, Vero, E5, or D14 cells were infected with d120 at a multiplicity of 1 pfu/cell and processed for immunofluorescence at 2 hpi (Fig 18C). White arrows in the merged panels indicate the juxtaposed ICP4 and PML foci. A d3-10 focus can clearly be

observed next to a PML focus in infected D14 cells indicating d3-10 localized to sites of viral genome deposition.

Chromatin immunopreciptation (ChIP) assays were performed to determine if d3-10 bound viral DNA. Vero or D14 cells were infected with d120, at multiplicity of 10 pfu/cell for 3.25 hours. KOS infected Vero cells were used as a positive control in this assay, while i13 or d120 infected Vero cells served as negative controls (284). The viral mutant i13 produces a full length ICP4 protein that is defective in binding to DNA due to a small insertion within the DNA binding domain (293). A monoclonal antibody that reacts with the C-terminus of ICP4, 58S (298), was used in the ChIP experiment, and the percentage of genomes containing ICP4 bound to the tk promoter was analyzed. The results from a representative assay are shown in Figure 18D. The ChIP assays were performed on three independent biological samples (Figure 18E). d3-10 and wtICP4 bound more genomes at the tk promoter than the negative controls in each independent sample. Taken together, the Western blot, immunofluorescence, and ChIP analyses demonstrated that D14 cells produced the d3-10 protein to comparable levels as a KOS infection and that the d3-10 protein localized to and bound viral DNA.

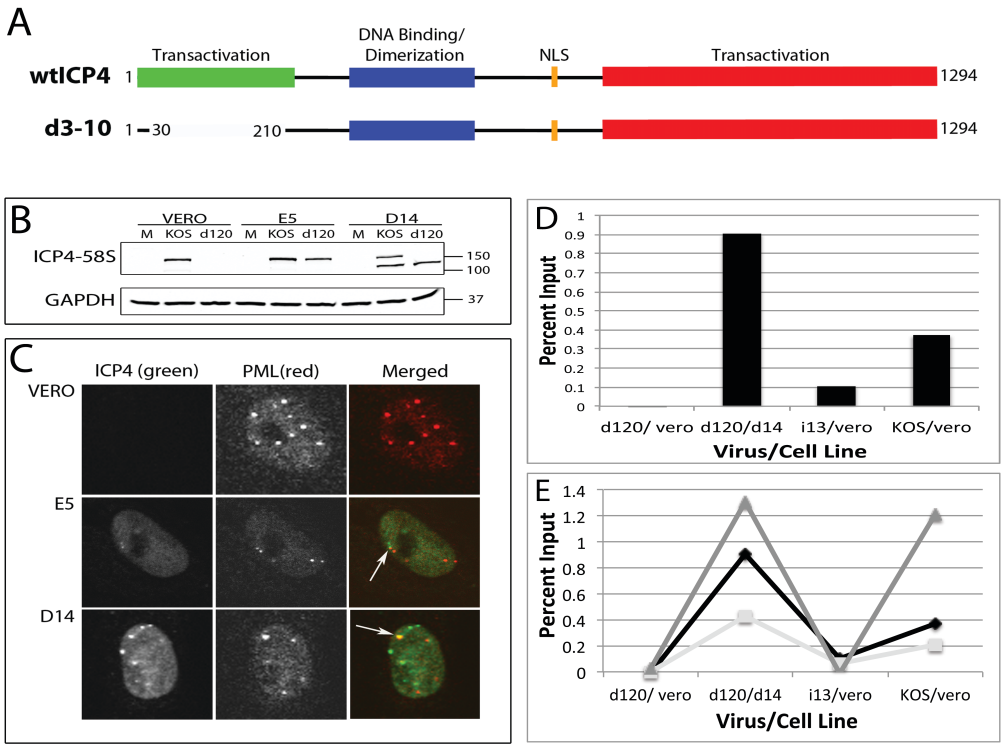


Figure 18: Characterization of d3-10 Expressing Cells.
(A) Schematic of the primary structure of ICP4 mutant molecules. (B) Western blot analysis of a d3-10 expressing cell line (D14 cells). Cells were mock infected or infected with KOS or the ICP4 null virus d120 at an MOI of 10. Infected whole cell extracts were subjected to western blot analyses using an antibody against the C-terminus of ICP4 (58S). (C) The localization of wtICP4 and d3-10 was compared to viral genome distribution (PML bodies) via immunofluorescence assays. (D) The ability of d3-10 to bind to viral DNA was assessed via chromatin immunoprecipitation. Data is presented as the percentage of input genomes bound by ICP4. (E) ChIP assays results from 3 independent biological samples.

ICP4-mediated Activation of E gene Expression Requires the N-Terminal Activation Domain.

The previous experiments established that the regions deleted in d3-10 did not affect the ability of the molecule to localize to or bind viral DNA, therefore the contributions of the deleted region to viral transcription and replication were addressed. To determine the contributions of amino acids 30-210 in viral replication, the ability of the d3-10 protein to complement an ICP4 null virus, d120, was examined. Vero, E5, or D14 cells were infected with d120 at a multiplicity of 10 pfu/cell. Viral lysates were harvested 24 hours post infection (hpi) and viral yields were

determined by plaque assay on the complementing E5 cell line. Figure 19A shows average viral yields from this assay. d120 infection of the wtICP4 producing E5 cells produced nearly 10^8 pfu/mL while infection of either Vero cells or D14 cells produced less than 10^4 pfu/mL. These data demonstrated that the d3-10 protein did not complement an ICP4 null virus, indicating a severe defect in ICP4 function.

The inability of the d3-10 protein to complement d120 suggested that d3-10 may be defective for the activation or repression of viral transcription. To investigate this, qRT-PCR was used to determine the relative quantities of tk and ICP4 mRNA present in d120 infected cell extracts at 4 hpi. Figure 19 B shows the percentage of tk mRNA quantities relative to that of E5 cells. Vero cells, which do not produce ICP4, expressed only 4.1% of the amount of tk mRNA as E5 cells. This represents transcription independent of ICP4 activation, and is similar to previously reported levels of tk RNA expression in the absence of functional ICP4 (141). Similarly, infection of D14 cells yielded only 4.6% of the amount of tk mRNA as E5 cells, indicating that the d3-10 protein was not capable of promoting E gene transcription.

ICP4 mRNA quantities from KOS, d120, or n12 infected Vero, E5, and D14 cells, were analyzed to determine whether amino acids 30-210 contributed to the repression of viral transcription. Upon activation of E genes, ICP4 binds to a site within its own promoter and prevents the activation of transcription, thus causing reduced accumulation of ICP4 mRNA (176, 177, 233). The viral mutant n12 expresses only the first 251 aa of the ICP4 protein, and as such, lacks the DNA binding domain rendering it defective for the repression of transcription (70). n12 infection of Vero cells yielded approximately 3 times more ICP4 mRNA than did infection of E5 cells, demonstrating the expected defect of n12 in repression (Figure 19 C). Conversely,

n12 infection of D14 cells yielded similar amounts of ICP4 mRNA as either KOS infected Vero cells or n12 infected E5 cells, indicating that the d3-10 protein retained repression activity.

While mRNA quantification data indicated amino acids 30-210 were dispensable for the repression functions but necessary for the activation functions of ICP4, these conclusions were drawn based on two representative genes. To obtain a better understanding of global expression patterns, protein synthesis profiles from d120 infected Vero, E5, and D14 cells were analyzed. Vero, E5, and D14 cells were infected at a multiplicity of 10 pfu/cell in methionine free medium for 3.5, 7.5, or 11.5 hours, at which point S^{35}-Methionine was added to cultures and allowed to incorporate into synthesizing polypeptides for 30 minutes. The polypeptides were separated by SDS-PAGE and visualized by autoradiography. Results of these assays showed reduced levels of IE protein synthesis in d120 infected D14 cells compared to Vero cells, despite equal amounts of cellular extract being loaded (Figure 19D; see ICP4, ICP0, ICP22, ICP27 and actin). Interestingly, D14 cells showed only IE protein synthesis; E and L protein expression was absent in D14 cells when compared to E5 cells (Figure 19D; see VP1/2, ICP8, gB, pgB, VP16). Taken together, the complementation, mRNA abundance, and protein expression data indicated that the amino acids 30-210 of the N-terminal transactivation domain are necessary for the activation of E gene transcription, but not IE gene repression.

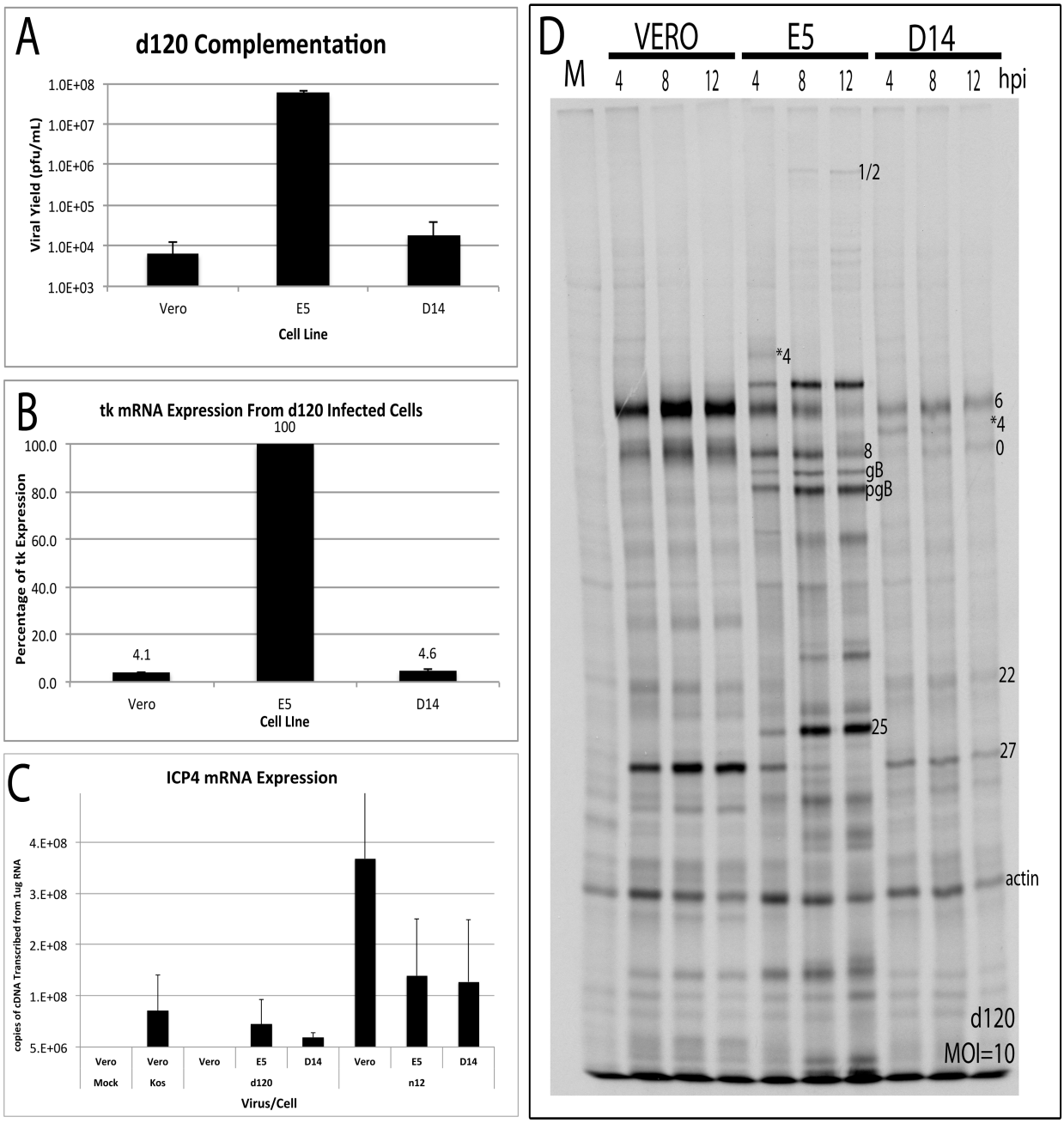


Figure 19: The N-terminus of ICP4 is Necessary For Trans-activation But Not Repression.
(A) Viral yield from d120 infected cells. Vero, E5, or D14 cells were infected at a multiplicity of 10 pfu/cell with the ICP4 null virus, d120. Viral lysates were collected 24 hpi, and viral yields were assessed by plaque assay on the E5 complementing cell line. (B) tk or (C) ICP4 mRNA expression levels in infected cells. RNA was harvested from Vero, E5, or D14 cells

infected with the indicated virus at a multiplicity of 10 pfu/cell for 4hrs. 1 ug of mRNA was reversed transcribed into cDNA, and tk and ICP4 mRNA was quantified using qRT-PCR. (B) Numbers are representative of the percentage of tk expression compared to that in wtICP4 expressing E5 cells. (D) Protein synthesis in d120 infected cells. Vero, E5, or D14 cells were infected at a multiplicity of 10 pfu/cell for 3.5, 7.5, or 11.5 hours with d120, or D14 cells were mock infected (M). Cells were incubated with 22 uCi ^{35}S-Methionine for 30 minutes. Labeled polypeptides were separated via SDS-PAGE and visualized via autoradiography. Protein identities are labeled.

The N-terminus of ICP4 is required for transcription complex formation

To address the basis for the reduced levels of E and L transcription in d3-10 expressing cells, the ability of PolII to bind to viral promoters was examined using a ChIP assay. Vero or D14 cells were infected with d120, i13, or KOS at a multiplicity of 10 pfu/cell for 3.25 hours. Immunoprecipitation was performed with an antibody directed against PolII and immunoprecipitated DNA was amplified with primers specific to the tk promoter. The results from the ChIP assays demonstrated that PolII bound larger percentages of genomes at the tk promoter in KOS infected cells than in d3-10 expressing cells. The percentage of genomes bound by PolII at the tk promoter in d3-10 expressing cells was similar to that observed in d120 or i13 infected Vero cells (Figure 20A). ChIP assays were performed in triplicate on independent biological samples and results are shown in Figure 20B. These data demonstrated that PolII is absent on a representative E promoter in d3-10 expressing cells, explaining the defects observed in E gene transcription (Figure 19).

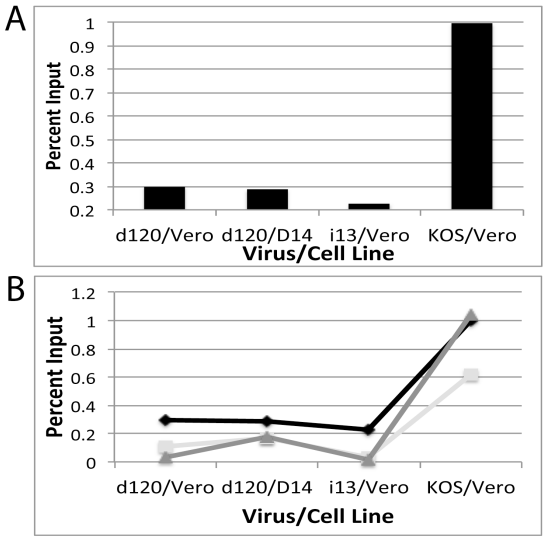


Figure 20: RNA PolII does not bind to Early promoters in d3-10 expressing cells.
(A) Percentage of genomes bound by PolII at the tk promoter. (B) Results from 3 independent biological samples.

The results from the ChIP assays demonstrated that the observed defects in E gene transcription in d3-10 expressing cells were the result of reduced PolII recruitment to viral promoters. TFIID and Mediator are largely involved in recruiting and stabilizing RNA PolII to TATA driven promoters and wtICP4 exists in complexes with these factors (198). To determine whether amino acids 30-210 were important for TFIID and Mediator interactions, immunofluorescence and affinity purification assays using d3-10 expressing cells were performed.

The cellular localization of TFIID and Mediator relative to d3-10 was investigated using immunofluorescence assays. Vero, E5, or D14 cells were infected with d120 at an MOI of 10 for 6 hr. Samples were processed for immunofluorescence with antibodies directed against ICP4 and TBP (Figure 21 A) or Med1 of Mediator (Figure 21 B). At 6 hpi wtICP4 was present in large replication compartments that included both TFIID and Mediator, consistent with previous findings (Figure 21 A & B) (198). Replication compartments are indicative of both high levels of transcription and DNA replication (375). d3-10 expressing cells were defective in E gene

transcription, and consequently would not replicate viral DNA, therefore, replication compartments did not form (Figure 21). Instead, d3-10 remained localized in small punctate foci within the nucleus, which are referred to as pre-replication compartments (Figure 21). TBP and d3-10 co-localized in some pre-replication compartments (Figure 21A, panel 3), suggesting that TBP and d3-10 formed a complex in infected cells. Interestingly, Mediator was not present in pre-replication compartments (Figure 21B, panel 3), suggesting a role for amino acids 30-210 in interactions with Mediator. To demonstrate that Mediator localized to pre-replication compartments with wtICP4, d120 infected E5 were also processed for immunofluorescence at 2 hpi, which coincides with pre-replication compartment formation in wild type infection. It is evident from the co-localization of wtICP4 and Mediator at 2 hpi, that Mediator was present in pre-replication compartments with wtICP4.

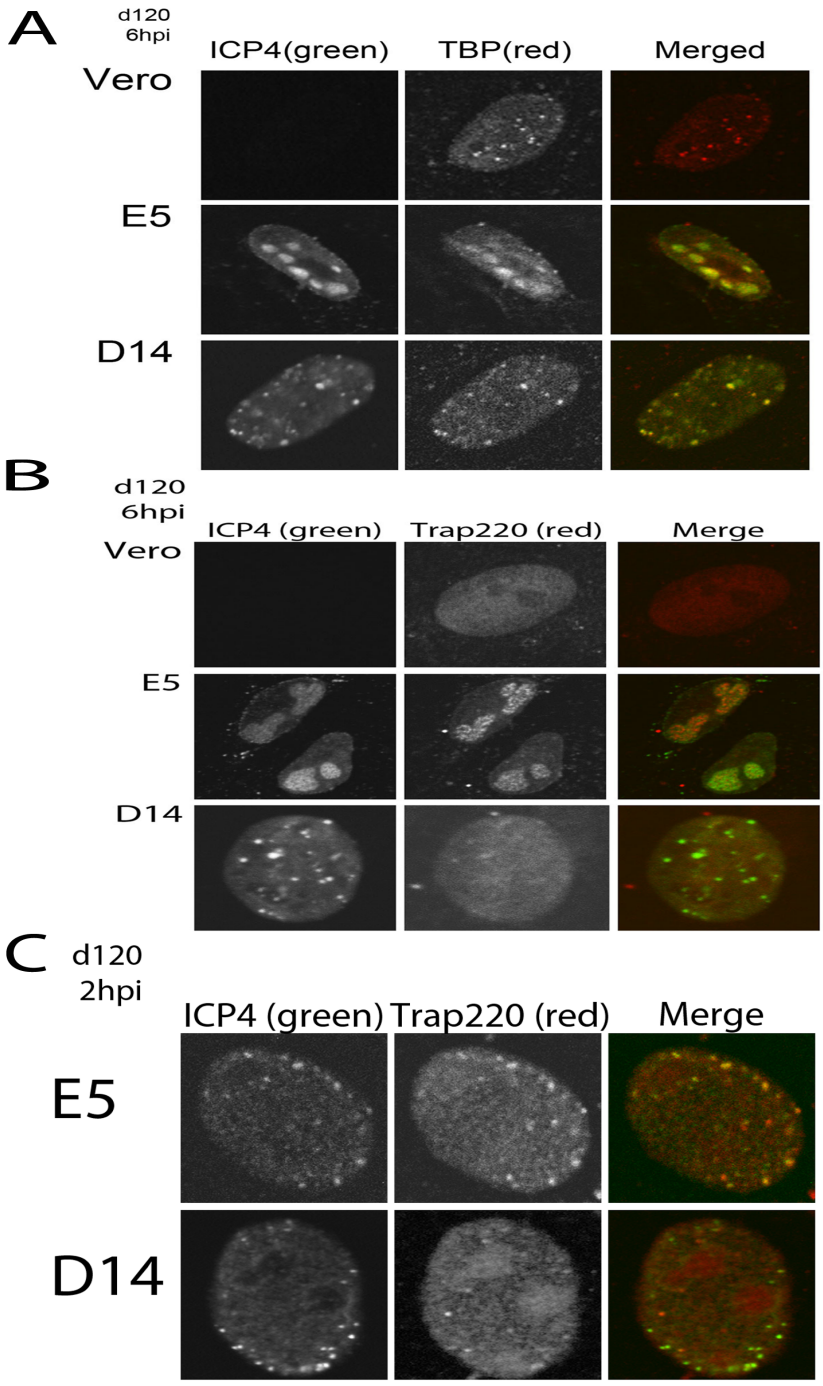


Figure 21: TBP is recruited to d3-10 complexes but Mediator is not.
Immunofluorescence assays of d120 infected cells. (A) ICP4 and TBP at 6 hpi. (B) ICP4 and Trap220 of Mediator at 6 hpi. (C). ICP4 and Trap220 at 2 hpi.

To assess whether d3-10 could form complexes with TFIID in infected cells, affinity purification of the d3-10 protein was performed. The d3-10 protein was genetically engineered

to contain a Tandem Affinity Purification (TAP) tag (Stratagene) at the N-terminus of the protein, and a stable cell line (I3) expressing the TAP-d3-10 under the ICP4 promoter was created. The TAP tag consists of a calmodulin binding peptide N-terminal to a streptavidin binding peptide. The TAP-tagged protein can then be purified using a streptavidin conjugated matrix, eluted with biotin, and further purified on a calmodulin-conjugated matrix if necessary. For the purposes of these studies purification using the streptavidin conjugated matrix was sufficient. As a control a TAP-tagged version of wtICP4 was constructed within the context of the viral genome and also purified over a streptavidin-conjugated matrix. Untagged versions of wtICP4 from KOS and d3-10 from D14 cells served as negative controls.

To identify whether TFIID was present d3-10 purified samples, immunoblots with antibodies directed against ICP4, TBP, or TAF1 were performed (Figure 22). The presence of strong immunoreactive bands at the correct molecular weights on the immunoblot of ICP4 indicated that TAP-ICP4 and TAP-d3-10 were isolated to significant quantities using this procedure (Figure 22; panel 1). Despite having approximately equal amounts of both TBP and TAF1 in all of the infected cell nuclear extracts, TBP and TAF1 were only present in the TAP-ICP4 samples, not the TAP-d3-10 (I3) samples (Figure 22; panels 2 & 3). This is evident by the absence of strong immunoreactive bands in the I3 pull down lanes of both the TBP and TAF1 immunoblot (Figure 22; panels 2 & 3). This result was perhaps surprising as TBP co-localized with d3-10 in infected cells. This may suggest that d3-10 can very weakly associate with TFIID in infected cells, but that this association is lost upon manipulation of the protein. Conversely, this may suggest that the d3-10 protein cannot associate with TFIID, and that the TFIID co-localization with d3-10 is ICP4 independent.

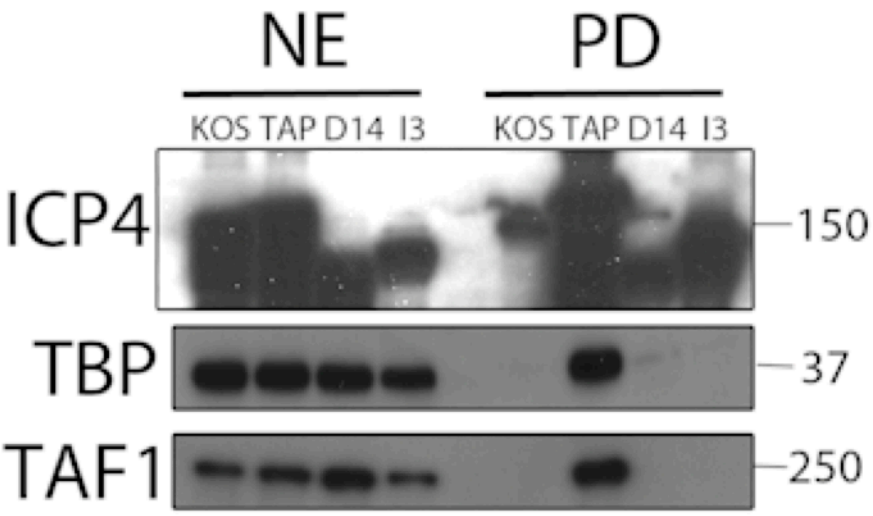


Figure 22: d3-10 does not co-purify with TFIID from infected cells.
KOS expresses wtICP4, TAP expresses TAP-ICP4, D14 cells express d3-10 and I3 cells express TAP-d3-10 upon infection with d120. ICP4 complexes were purified and the presence of TBP and TAF1 were determined by immunoblot for ICP4, TBP, and TAF1. Nuclear extracts (NE) and pull downs (PD) were used.

Intragenic Complementation of ICP4

ICP4 mutants can function together to complement the defects observed within each individual mutant (293, 294). For example, a mutant defective in DNA binding (d2) can be paired with a mutant defective for transactivation (n208) to form a functional molecule (294). The formation of a functional molecule can be attributed to the dimerization properties of ICP4. In this respect, it can be hypothesized that d3-10, which is missing the N-terminal transactivation domain, and n208, which is missing the C-terminal transactivation domain could form a functional heterodimer and complement the defects observed in gene expression (69, 70).

To investigate this a complementation assay was performed. Vero, E5, and D14 cells were infected with n208 at a multiplicity of 5 pfu/cell, infected cell lysates were harvested at 24 hpi, and viral yields were determined by titration on the complementing E5 cell line. In accordance with previously published data, yields from n208 infected Vero cells were reduced

by more than 1.5 log when compared yields for the complementing E5 cell line (Figure 23 A) (70, 340). Interestingly, n208 infection of D14 cells yielded similar levels of viral progeny as infection of E5 cells (slightly under 10^7 pfu). Therefore, n208 and d3-10 complement each other.

n208 has previously been shown to be defective in forming replication compartments, a hallmark of DNA replication and L gene expression (294). As shown in Figure 21, d3-10 expression alone did not result in the formation of replication compartments. Instead, d3-10 formed small aggregates in the nucleus, consistent with pre-replication compartments (169, 207). It was hypothesized that n208 infection of D14 cells would result in the production of viral replication compartments because n208 and d3-10 complement at the level of viral replication. To more closely investigate the nuclear distribution of n208 in infected D14 cells, immunofluorescence assays were performed. D14 cells were infected with n208 for 6 hours and samples were processed for immunofluorescence (Figure 23 B). d3-10 was detected with the C-terminal ICP4 antibody 58S while n208 was detected with the N-terminal antibody N15. Immunofluorescence assays demonstrated that, consistent with previously published data, n208 expression in Vero cells yielded a robust nuclear stain, with no discernable foci or replication compartments (Figure 23 B panel 1) (294). Conversely, n208 infection of D14 cells yielded cells with large nuclear replication compartments (Figure 23 B single cell), indicative of increased viral transcription and viral DNA replication. The replication compartments formed in n208 infected d3-10 cells were similar in morphology to previously published work and would be consistent with the increased viral yields seen in Figure 23 A (169, 375).

To determine whether d3-10 and n208 complementation was the result of the formation of functional heterodimers, a TAP-tagged version of n208 was utilized. To establish that d3-10 and n208 could form a heterodimer, D14 cells were infected with n208 or TAP-n208, TAP-n208

containing complexes were affinity purified, and the presence of d3-10 in these complexes was assessed via immunoblotting. Figure 23 C is an immunoblot for n208 (N15) and d3-10 (58S). As a control, n208 infected Vero and E5 cells were also utilized in this assay. Affinity purification of TAP-n208 from E5 cells yielded two bands in the pull down (PD), one corresponding to n208 (red) and one corresponding to wtICP4 (yellow). This indicated that wtICP4 and TAP-n208 could form heterodimers and validated the use of this assay. Additionally, a portion of wtICP4 (green) remained in the flow through (FT), and corresponded to wtICP4 homodimers. Affinity purification of TAP-n208 from infected D14 cells resulted in the presence of two bands in the PD (both d3-10 [green] and n208 [red]; lane 12), and only one band in the flow through (lane 11) corresponding to the d3-10 protein that did not form heterodimers with TAP-n208.

Affinity purification data suggested that the defects observed with the d3-10 mutant were a result of ablated interactions with TFIID. To determine whether the formation of an n208-d3-10 heterodimer reconstituted this function, TAPd3-10 expressing cells were infected with n208 and complexes were affinity purified over a streptavidin matrix (Figure 23 D). This provided for isolation of proteins in complex with only TAPd3-10 homodimers and TAPd3-10-n208 heterodimers. As controls, Vero cells were infected with either KOS or TAP, which contains a TAP-tagged full length ICP4, and D14 and I3 cells were infected with d120. The presence of an immunoreactive band in the TAP PD but not I3 PD lanes in an immunoblot for TBP indicated that TBP was isolated with TAP-ICP4, but not with TAPd3-10 (Figure 23 D, lanes 4 and 12). Significantly, purification of TAPd3-10-n208 heterodimers resulted in isolation of TBP (Figure 23D, lane 8). Taken together, these data support a model in which the complementation of n208 and d3-10 occurs through the ability of the two proteins to form a more functional dimer.

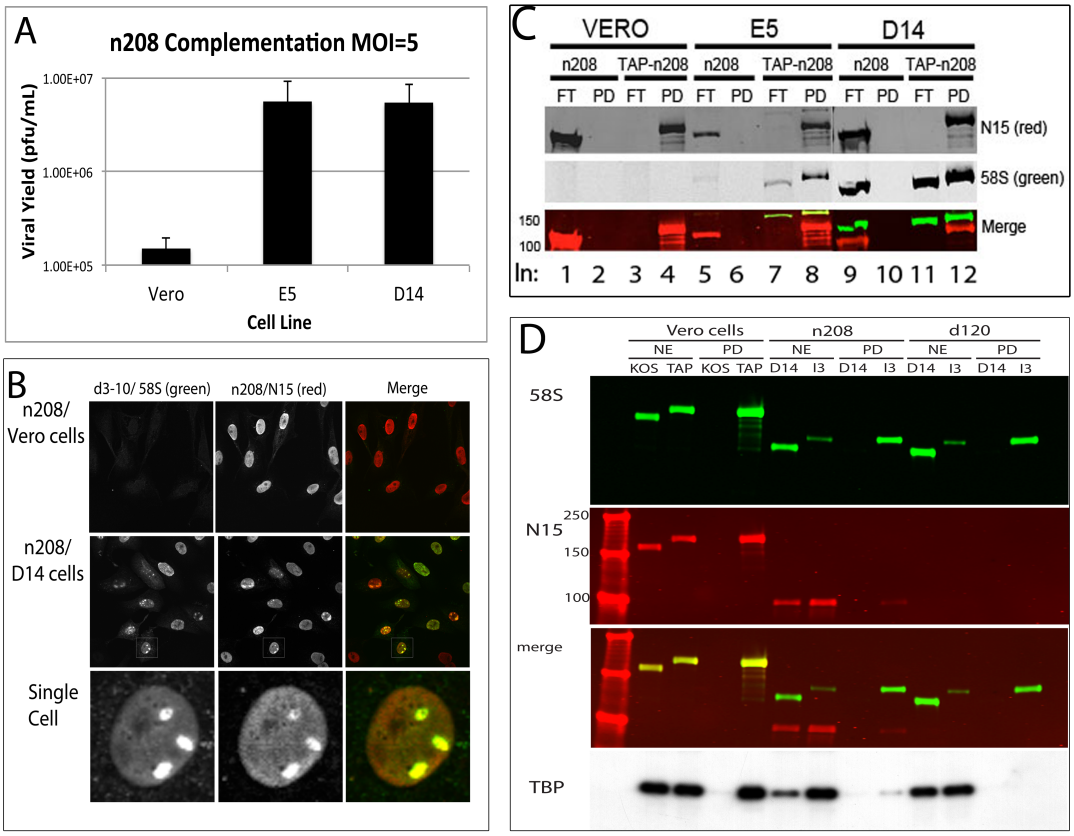


Figure 23: Intragenic complementation of ICP4.
(A) Complementation assay from n208 infection of Vero, E5, and D14 cells. (B) Immunofluorescence with antibodies specific for d3-10 (58S; green) and n208 (N15; red). (C) Western blot analysis of TAP purified ICP4 complexes. Both the flow through (FT) and pull down (PD) were analyzed. Antibodies for d3-10 (58S; green) and n208 (N15; red) were used for the immunoblot. (D) Western blot analyses of affinity purified d3-10-n208 heterodimers.

3.5 DICUSSION

ICP4 has two discontinuous regions that are involved in the transactivation of transcription, one N-terminal and one C-terminal to the DNA binding domain. Previous studies demonstrated that the N-terminal transactivation domain is sufficient for viral growth (70). This study investigated the requirement of the N-terminal transactivation domain to the activities of

ICP4 with respect to viral gene transcription, and the possible sufficiency of the C-terminal transactivation domain. The data demonstrated that the N-terminus of ICP4 was necessary for ICP4 mediated transcription, as early gene expression was ablated when the d3-10 mutant protein was expressed. The C-terminus and the DNA binding domains were unable to activate transcription, however, supplying the N-terminus in trans could complement this defect. The proposed causes for the transcription defects observed are discussed below.

It has been previously demonstrated that ICP4 interacts with a variety of transcription factors including TFIID and Mediator to stabilize transcription initiation machinery on HSV promoters (41, 117, 198, 284, 373). The data presented herein demonstrated that amino acids 30-210 of the N-terminal transactivation domain were required for the activation of viral transcription through the activities of these regions in forming complexes with TFIID and Mediator. The N-terminus of ICP4 contains several stretches of 10-20 highly conserved amino acids that individually contribute to viral growth or transcription (340). Here we show that deletion of all of these regions, in addition to the degenerate, disordered regions flanking them resulted in the ablation of E transcription. This suggests that the N-terminus may form a tertiary structure that enhances the binding of transcription complexes, such as TFIID and Mediator.

The N-terminus of ICP4 contains stretches of amino acids that are conserved in other ICP4 analogs and are predicted to be relatively disordered, respectively (340). The flexibility of the disordered regions and the specificity of the conserved regions likely allow ICP4 to interact with multiple surfaces of transcription complexes. A key example of this is the interaction with and stabilization of TFIID. The C-terminus of ICP4 has been shown to interact with TAF1 while the N-terminal 774 amino acids were sufficient for interactions with TBP, both components of TFIID (41, 303). Deletion of amino acids 81-96 in addition to the C-terminal transactivation

domain eliminated ICP4 mediated viral transcription, perhaps suggesting that both of these regions stabilize interactions with TFIID (340). Interestingly, deletion of amino acids 30-210, in the presence of the C-terminus, completely abolished interactions with TFIID, while deletion of amino acids 81-96 alone presumably did not. The inability of d3-10 to form complexes with TFIID is probably the result of deletion of multiple TFIID interfaces within the N-terminus, without which the C-terminal interfaces may be too weak to compensate, suggesting that the activities of the N-terminal transactivation domain are augmented by functions of the C-terminus. Alternatively, amino acids in the C-terminus and N-terminus affect the conformation of each other.

ICP4 exists as a dimer within infected cell nuclei (223). Previous studies have shown that individual defects in ICP4 can be complemented by the formation of a heterodimer. For example, coinfection of a virus encoding an ICP4 molecule deficient in DNA binding activities, d2, and the C-terminal truncation mutant, n208 yielded a heterodimer that was capable of binding to DNA, suggesting that one functional domain within a dimer was sufficient for function (294). Additional electromobility shift and supershift assays revealed that the conformation of the C-terminal transactivation domain of wtICP4 was negatively altered in a heterodimer consisting of wtICP4 and the mutant X25, in which both the N-terminal and C-terminal transactivation domains were deleted (296). This strongly suggested that the structure of one molecule within the dimer directly affected that of the other. The functional contributions of the N-terminal and C-terminal transactivation domains to the formation of a functional dimer had not been established. Infection of d3-10 expressing cells with n208 yielded complementation at the level of viral production, but more importantly reconstituted the ability of the d3-10 molecule to interact with TFIID. This demonstrated not only that one functional N-terminal and

C-terminal transactivation domain are sufficient for viral replication, but also that the N-terminus and C-terminus of ICP4 cooperate to mediate the activities of ICP4. Additionally, this implies that the N-terminal and C-terminal transactivation domains of opposite molecules within the dimer are capable of functioning together, perhaps suggesting a conformation in which they are in close proximity within the dimer.

The data presented here demonstrated that the N-terminal transactivation domain is absolutely necessary for ICP4 mediated transcription, yet this region remains very poorly conserved amongst alphaherpesviruses. Interestingly, while ICP4 and the VZV homolog IE62 have minimal sequence homology within the amino terminus, IE62 can partially complement for the absence of ICP4 (98). This suggests that these viruses evolved to contain similar functions, without maintaining sequence homology. Interestingly, the Ruyechan lab has shown that the N-terminus of IE62 specifies for interactions with the Mediator complex, a function that, as shown here, at least partially resides within the N-terminus of ICP4 (366, 368).

The work presented herein suggests that ICP4 is a structurally complex molecule functioning to provide an interface between multiple viral transcription complexes and viral promoters. The ability of ICP4 to interact with components of the general transcription machinery is the direct result of the structural conformations adopted by the dimeric complex. The functions provided by ICP4 are integral for viral replication, and as such ICP4 contains redundancies throughout the molecule to ensure its function within infected cells.

3.6 ACKNOWLEDGEMENTS

The I3 cell line was constructed by the graduate student, Avaraham Bayer.

4.0 HSV-1 ICP4 PROVIDES A PLATFORM FOR THE ASSEMBLY OF TRANSCRIPTION COMPLEXES

4.1 ABSTRACT

Herpes Simplex Virus utilizes general transcription factors (GTFs), activators, and RNA polymerase II from the host cell to transcribe its genome. In addition to cellular factors, the viral encoded transactivator, ICP4, promotes the transcription of viral genes by stabilizing transcription machinery onto viral promoters. Previously, our lab has shown that ICP4 forms complexes with TFIID and Mediator in infected cells. Here, we use affinity purification of ICP4 and western blot analyses to show that there is an ordered recruitment of these factors into ICP4 containing complexes. TFIID is present at higher levels in ICP4 containing complexes earlier in infection than mediator, indicating a temporal regulation of transcription factor and activator recruitment. Additionally, we can show that the N-terminal 774 amino acids of ICP4 are sufficient to form a complex with TFIID and mediator, although these interactions are not as strong as with wtICP4. Perhaps most strikingly, we demonstrate that ICP4 is not only isolated with components involved in transcription initiation, but also chromatin modification, transcription elongation, and potentially mRNA processing. Our data show that ICP4 can be isolated with the chromatin remodeling factors SWI/SNF, NURD, and the Ino80 complex, indicating that ICP4 may have a significant role in chromatin remodeling. Additionally, the

elongation related factors TFIIH and the elongin/SIII complex were isolated in ICP4 containing complexes. Together these data indicate that ICP4 plays a significant role in mediating HSV transcription, and that ICP4 may act as hub for cellular transcription factor recruitment, affecting multiple steps in the formation of transcription complexes.

4.2 INTRODUCTION

ICP4 is an essential viral protein that enhances RNA PolII transcription of viral genes. The activities of ICP4 are defined by two transactivation domains separated by a DNA binding domain and a nuclear localization sequence (70, 253, 254, 295). The DNA binding activities are essential, but not sufficient, for both the transactivation and repression functions of the molecule (117, 253, 254, 295). The N-terminal and C-terminal transactivation domains are also necessary for viral transcription, although to different extents (340). We have recently shown that the N-terminal transactivation domain of ICP4 is necessary for E, and as a consequence, L, gene transcription (Chapters 2 and 3). The C-terminal 500 amino acids are necessary for abundant L gene expression; without the C-terminus viral growth is attenuated by approximately 100 fold (69, 70, 295, 340). The defect observed in viral growth in the absence of either the N-terminus (d3-10) or C-terminus (n208) may be an effect of diminished interactions between ICP4 and cellular transcription machinery (Chapter 3).

ICP4 has been shown to interact with a number of components of the cellular transcription machinery. It was demonstrated that ICP4 directly interacts with the TAF1 and TBP components of TFIID (41, 119, 303, 373). These interactions have been mapped to the C- and N-termini of ICP4, respectively, suggesting that multiple surfaces of ICP4 contact the TFIID

complex (41, 182, 373). The interaction with multiple subunits of TFIID on differing surfaces of ICP4 underscores the importance of this interaction. Additionally, it is probable that only a subset of these interactions are necessary for stabilization of TFIID on promoters, as it has been shown that ICP4 contains redundant functions within the N- and C-termini (340). More recently, ICP4 has been shown to co-purify and localize with Mediator in infected cells (198). The structural requirements for interactions with Mediator have not yet been clearly defined, although the N-terminus appears to be important (Chapters 2 and 3).

The current model for viral transcription suggests that the transcription complexes formed on the promoters of E and L genes partially defines the temporal kinetics of viral transcription. Interestingly, it has been shown that TAF1 is necessary for L gene expression (73). The C-terminus of ICP4 is necessary for both interactions with TAF1 and L gene transcription (41, 70). These data suggest a potential mechanism of activation of L genes that relies on the ability of ICP4 to interact with TAF1 to drive Inr dependent transcription of L genes (163). Likewise, it has been shown that TFIIA is critical for the activation of early, but not L genes, partially due to its ability to stabilize TBP, and hence TFIID, to TATA boxes (372, 373), providing yet another means for regulation of temporal transcription. Thus, the tertiary structure of ICP4 defines the transcription complexes with which it can interact, which in turn, defines temporal transcription from the viral genome.

The goals of this study were to investigate the structural and temporal basis for ICP4 mediated interactions with components of the transcription machinery and to isolate novel ICP4 containing complexes. To isolate ICP4 containing protein complexes, wtICP4, and the C-terminal truncation mutant n208, were genetically engineered within the viral genome to contain a tandem affinity purification (TAP) tag. ICP4 containing complexes were purified using a

streptavidin conjugated matrix and associated proteins were identified using LC-MS-MS and confirmed via western blot analysis. Data indicates that ICP4 interacts with TFIID prior to the Mediator complex and that the N-terminal 774 amino acids are sufficient, but sub-optimal, for these interactions. Additionally, ICP4 co-purifies with components involved in all steps of transcription including chromatin remodeling, initiation, elongation, and possibly mRNA processing and export. Together the data suggest that ICP4 may play a significant role in regulating transcription at multiple levels.

4.3 MATERIALS AND METHODS

Cells and Viruses. Vero cells (African Green Monkey Kidney cells) were cultured in Dulbecco's Modified Eagle Medium with 5% FBS and maintained as suggested by ATCC. E5 cells express complementing levels of wtICP4, and have been described previously (69). The viruses wild type strain KOS and n208 (70) have been previously described. TAP-KOS and TAP-n208 were generated for this study via marker transfer with the TAP-pK1-2 and TAP-pn7 plasmids (described below), plaque isolated a minimum of three times, and the genotypes were confirmed via southern blot and Sanger sequencing.

Recombinant Plasmids. TAP-ICP4 plasmids were constructed such that the TAP sequence would be in frame and 5' to the ICP4 start site. wtICP4 and n208 plasmids, pk1-2 and pn7 respectively, were utilized to construct the TAP-pk1-2 and TAP-pn7 plasmids via RecET recombination system, as described previously (327, 328). Briefly, primers were designed such that they would amplify the TAP sequence and a kanamycin resistance cassette from the pTAP-

KANinv plasmid (a gift of Dr. Paul Kinchington, University of Pittsburgh) and contained homology to ICP4 –60 to –20 and -7 to +31, respective to the ICP4 transcription start site (tss). The sequences for the primers used to construct the mutants were 5'CTACCGTGCTACGTCCGCCGTCGCAGGCCGTATCCCCGGAGGTTTAGTGAACCGTC AGATCCGCTAG3' and 5'GAGCCGGGGCGCTGCTTGTTCTCCGACGCCATCGCCGATGGATCCGCCCGGGCGC CCAGCTTGCAG3'. Primers were synthesized and gel purified by IDT (Coralville, IA). Purified PCR product was transformed into GS1783 cells (a gift of Greg Smith, Northwestern University) containing pk1-2 or pn7. The appropriate colonies were selected for their resistance to kanamycin and screened via restriction endonuclease and agarose gel analysis. Colonies containing the appropriate insert were induced with 1% arabinose at 42° to allow for production of the SceI enzyme and recombination to occur. The resultant colonies were screened for sensitivity to kanamycin and possession of the appropriate allele via restriction endonuclease and agarose gel analysis.

The m90-GST plasmid was constructed using standard cloning techniques to insert the m90 region downstream of a sequence encoding GST in the pGEX-2T plasmid (gift of Dr. Martin Schmidt, University of Pittsburgh). Briefly, the m90 region of ICP4 (+224-+307 bp relative to the transcription start site) was amplified via PCR using primers with homology to ICP4 and such that EcoRI and BamHI endonuclease recognition sites flanked the m90 region. The primers, 5'AGCTTTGGGTGAACAGGATCCGACGCCGTCTCGCCGCGACAGCTG3' and 5'AAATTGGTAACTGCTGAATTCTGCGTCGGGATCGTCCGGACGGCCTC3', were synthesized by IDT (Coralville, IA). Amplified, PCR product and plasmid pGEX-2T were subjected to restriction endonuclease digestion followed by ligation using T4 DNA Ligase

according to the manufacturers suggestions (NEB). Ligated products were transformed into DH5α *E.coli*. Colonies were selected based on resistance to carbanecillin and the genotype was verified via restriction endonuclease and agarose gel analysis and Sanger sequencing (Genewiz, South Plainfield, NJ).

Streptavidin Affinity Purification: $2x10^8$ Vero cells were infected at a multiplicity of 10 pfu/cell with KOS, TAP-ICP4, n208, or TAP-n208 for 3, 6, or 12 hours at 37° prior to harvesting. Cells were rinsed twice with TBS containing 0.1 mM *N-p*-tosyl-L-lysine chloromethyl ketone (TLCK) and scraped into 50 mL TBS + 0.1 mM TLCK. Cell pellets were collected via centrifugation at 5 K for 5 min in the Sorvell RC5 centrifuge. Nuclei were collected and proteins were extracted in 0.4 M KCl as described previously (76, 198). Nuclear extracts were diluted in half in 2 X low salt streptavidin binding buffer (2 X LS SBB) (20 mM Hepes KOH pH 7.9, 20 mM KCl, 4 mM EDTA, 0.2% NP40), treated with 300 U benzonase nuclease (Novagen) and incubated with 500 µL streptavidin conjugated agarose beads (Pierce) overnight at 4° with end over end rotation. Beads were washed 4 times in 10 mL of streptavidin binding buffer (SBB) (20 mM Hepes KOH pH 7.9, 200 mM KCl, 2 mM EDTA, 0.1% NP40). Samples were eluted from the beads using 500 µL streptavidin elution buffer (SBB + 2 mM Biotin). Two elutions, each 30 min, and one elution, overnight, were combined and either concentrated in 7 mL / 9 kDa iCon concentrator (Pierce) and used for western blot analyses, or sent to MS Bioworks (Ann Arbor, MI) for TCA precipitation, in gel digestion, and LC-MS-MS analysis.

GST Purification. DH5α cells containing either m90pGEX-2T or pGEX-2T were grown in Luria Broth at 37° to an OD_{600} of 0.5. Protein expression was induced with 0.1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) for 2hrs. Cells were harvested, washed in 1 X PBS (137 mM NaCl, 2.68 mM KCl, 10.14 mM $NaHPO_4$, 1.76 mM KH_2PO_4, pH 7.4), and proteins were extracted via sonication for 12-20 second pulses at a 50% output (Misonix sonicator 3000 with microtip). GST tagged protein extracts were incubated with 150 µL of 50% slurry of equilibrated glutathione sepharose 4B beads (GE#17-0756-01) in PBS. Unbound proteins were washed away with five 5 minute washes in PBS and one 5 minute wash in Low Salt Buffer (100 mM KCl, 1.5 mM MgCl2, 20% Glycerol, 20 mM Hepes) in the presence of protease inhibitor cocktail (Roche). Nuclear extracts from uninfected HeLa cells (372) were incubated with the m90GST and GST peptides immobilized on the glutathione conjugated beads. Unbound proteins were removed with 6 washes for a total of 30minutes in Low Salt Buffer + 0.1 mM TLCK. Proteins were eluted in 500 µL Elution Buffer (10 mM Glutathione, 200 mM NaCl, 50 mM Tris-HCl, pH 8.8) 3 times for 30 minutes each at 4°. Eluates were concentrated using iCon 7 ml / 9 KDa Concentrators (Pierce). Samples were analyzed via silver stain, Western blot, and LC-MS-MS.

LC-MS-MS Analysis. Elution samples were sent to MS Bioworks (Ann Arbor, MI) for processing. Briefly, protein samples were precipitate with TCA as previously described(148). Proteins were separated on a 10% Bis-Tris Novex mini-gel using a MES based system, stained with Coomassie stain, and the gel was separated into 10 equal sized slices. Gel fragments were processed using the ProGest (DigiLab) robot. The gel slices were washed with 25 mM ammonium bicarbonate followed by acetonitrile. The samples were reduced with 10 mM

Dithiothreitol (DTT) at 60° followed by alkylation with 50 mM iodoacetamide at room temperature. The protein samples were digested with trypsin (Promega) for 4 hours at 37° which was quenched with formic acid. The supernatant was then loaded onto a Waters NanoAcquity HPLC system interfaced to an Orbitrap Velos Pro (Thermo-Fisher). Data was analyzed against the *Rhesus macaque* and *HHV*-1 proteomes using the MASCOT program and validated using Scaffold software.

Electrophoresis, Silver Stain, and Western Blot. SDS polyacrylamide gel electrophoresis was carried out as previously described (340) with the following modifications. Samples were loaded onto either a 7.5% (ICP4, Med1, TAF1, Brm1), 10% (p62), any kD (TBP, Elongin B), or 4-15% (Silver stain) Tris-HCl mini protein gel containing a 4% stacking gel (Biorad). For silver staining, gels were processed according to the manufacturers directions (Pierce). For western blotting, proteins were transferred to either polyvinylidine fluoride (PVDF; Amersham) for chemi-luminescent detection or nitrocellulose membranes for infrared detection as previously described (340). Membranes were probed with the following antibodies in TBS-T containing 1% milk; N15, polyclonal rabbit serum, for ICP4 (1:500), Trap220 for Mediator1 (sc8998-X) (1:250), TAF1 (sc735) (1:500), TBP (233-R Covance) (1:250), p62 of TFIIH (sc292-X) (1:250), Brahma of SWI/SNF (sc6450) (1:250), and Elongin B of SIII (sc11447) (1:250). Membranes were processed as described previously (340) with the following modifications. For chemiluminescent detection, blots were probed with secondary HRP conjugated goat anti-rabbit, goat anti-mouse, or donkey anti-goat antibodies (Promega) at a 1:5000 dilution for 30 minutes. Membranes were then exposed to ECL detection agent (Amersham) according to the manufacturers instructions, and exposed directly to film (Hyperfilm; Amersham). For infrared

detection, membranes were incubated with IRdye conjugated donkey anti-rabbit or goat anti-mouse secondary antibodies at a 1:10,000 dilution for 30 minutes. Infrared signals were detected using the LI-COR Odyssey infrared imager.

4.4 RESULTS

The goal of this work was to investigate the complexity of ICP4 mediated interactions throughout viral infection. In this regard, the temporal kinetics and genetic requirements of said interactions were investigated. The identification of novel protein complexes that play an integral role in the regulation of ICP4 activated transcription was also a goal. To address these goals, affinity purification of wtICP4 and the C-terminal truncation mutant n208, followed by proteomic analyses were preformed.

Construction and characterization of TAP-ICP4 and TAP-n208

To investigate the protein components of ICP4 complexes, an affinity purification approach was taken. To affinity purify ICP4 containing complexes a Tandem Affinity Purification (TAP) tag (Stratagene) was placed at the amino terminus of the ICP4 protein. The TAP tag consists of a Calmodulin Binding Peptide (CBP) N-terminal to a Streptavidin Binding Peptide. Using the RecET system of recombination that has been previously described (327, 328), the sequence for the TAP tag was inserted into an ICP4 or n208 coding plasmid N-terminal to the ICP4 start site and such that it would be in frame with ICP4. The TAP tagged ICP4 or n208 containing plasmids were then used in a marker transfer assay to insert the TAP tag into both ICP4 loci within the viral genome. The genotypes of the resultant viruses were verified via Southern blot analyses and protein expression was verified via Western blot analyses (data not

shown). Both viruses had the correct genotype and expressed similar quantities of ICP4 with a decreased mobility when compared to the parent virus, corresponding to the addition of the TAP tag.

Characterization of ICP4 containing complexes.

To determine the composition of the protein complexes with which ICP4 interacts during infection, TAP-ICP4 or TAP-n208 were affinity purified using a streptavidin conjugated matrix, unbound proteins were washed away, and purified complexes were eluted three separate times using 2 mM biotin. To initially examine the composition of the ICP4 containing complexes, proteins were separated via electrophoresis and visualized using silver stain. Figure 24 is a representative silver stain for KOS and TAP-ICP4 purified at 6 hpi. It was evident from the presence of a strong band below 200kDa in the TAP-ICP4 elutions E1, E2, and E3 lanes that TAP-ICP4 was highly enriched using this procedure. Additionally, numerous other protein bands were visualized in the TAP-ICP4 elutions that were either not present or underrepresented in the KOS samples, indicative of the purification of ICP4 interacting proteins.

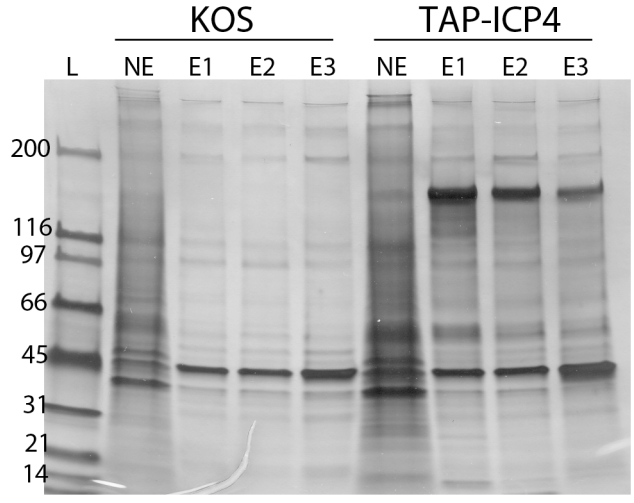


Figure 24: Silver stain analysis of ICP4 purification.

Vero cells were infected at a multiplicity of 10, nuclear extracts were collected 6 hpi, and ICP4 containing complexes were purified over streptavidin conjugated agarose beads. (NE: nuclear extract; E1-E3: Elutions 1-3). Note that each lane contains 1/20 of the eluted sample.

Mass Spectrometry.

ICP4 differentially regulates E and L transcription. Because E and L genes differ in their requirements for activation and because the core promoter elements are different between E and L genes, the temporal affect of ICP4 mediated protein interactions were investigated. KOS and TAP-ICP4 infected samples were collected at 3, 6, and 12 hpi and ICP4 containing complexes were affinity purified. Additionally, the viral mutants n208 and TAP-n208, were also purified to investigate the genetic requirements for ICP4 mediated protein interactions.

To identify the protein composition of each TAP sample, the three eluted fractions were combined, precipitated, and liquid chromatography coupled with tandem mass spectrometry (LC-MS-MS) analyses were performed. As a control, all experiments were performed with viruses containing untagged versions of ICP4 and n208 and were included for LC-MS-MS analyses. Proteins containing the peptides identified via LC-MS-MS were identified using the MASCOT search engine and verified using Scaffold. Search results were filtered to contain proteins that were at least 4x more abundant in the TAP-tagged samples than in the untagged samples and had a minimum of 5 spectral counts. Based on these criteria, 82, 141, 215, and 188 proteins were identified in the TAP-ICP4 3, 6, and 12 hpi, and TAP-n208 samples respectively.

ICP4 mediated TFIID interactions precede those of Mediator.

Using a similar tandem affinity purification approach, our lab has previously demonstrated that TFIID and Mediator can be co-purified with ICP4 (198). LC-MS-MS analyses from the experiments herein confirmed these findings. The percent coverage of each TAF and TBP from TFIID identified by LC-MS-MS is listed in Table 1. The data indicated that

ICP4 and TFIID formed complexes as early as 3 hpi and that these interactions are retained throughout infection. Presumably, the spectral counts correlate with the quantity of protein purified; higher spectral counts equate to a greater amount of that protein in the sample. Under this presumption, TAFs 4, 5, 6, and 9 were present in the greatest amounts in purified ICP4 samples. Interestingly, these TAFs, along with TAF12 have been suggested to be the "core" components of TFIID and are thought to initiate the assembly of holoTFIID (356). The abundance of these TAFs in affinity purified samples is either the result of ICP4 directly contacting a core component of TFIID or is reflective of the stoichiometry of holo-TFIID (285).

	TAP 3h		TAP 6h		TAP 12h	
	SpC	% Coverage	SpC	% Coverage	SpC	% Coverage
TAF1	137	38	100	28	51	21
TAF2	92	44	67	34	41	26
TAF3	28	25	8	6.2	16	17
TAF4	69	57	43	50	33	49
TAF5	132	69	67	55	76	48
TAF6	96	65	76	55	51	58
TAF7	15	32	7	21	6	17
TAF8	26	54	11	22	8	39
TAF9	27	65	12	31	0	0
TAF10	10	22	8	25	7	22
TAF11	12	33	0	5.2	0	0
TAF12	24	31	14	27	19	30
TBP	25	22	17	28	10	12

Table 1: Isolation of TFIID.
TFIID components isolated with TAP-ICP4 at 3, 6, and 12 hpi represented in spectral counts (SpC) and percent coverage (% Coverage).

In addition to isolating TBP and 12 TAFs from TFIID, 18 of the approximately 30 subunits of the Meditator complex were purified with ICP4 containing complexes (Table 2). Subunits from each module of Mediator (head, middle, tail, and kinase) were isolated in ICP4

containing complexes (Table 2). Interestingly, at 3 hpi only 5 subunits of the total 18 identified were identified by LC-MS-MS. The percent coverage of the 5 subunits identified at 3 hpi was relatively low, ranging from 1.4 - 15%, compared to at 6 hpi where the percent coverage of those same 5 subunits ranged from 17 - 29%. This indicates that ICP4 mediated interactions with Mediator were not present initially in infection as they appeared to be with TFIID (Table 1), implying that ICP4-TFIID interactions preceded those of Mediator.

	TAP 3h		TAP 6h		TAP 12h	
	SpC	% Coverage	SpC	% Coverage	SpC	% Coverage
MED1	0	0	23	17	8	3.2
MED4	0	0	8	30	0	16
MED6	0	0	9	22	0	18
MED7	0	0	8	25	0	0
MED12	11	4.7	58	17	31	15
MED13	0	0	23	9.3	7	3.9
MED14	0	0	30	18	25	8.8
MED15	0	0	8	11	5	7.7
MED16	3	3.5	22	18	18	19
MED17	9	15	16	29	18	23
MED18	0	0	7	18	10	25
MED20	7	15	14	28	14	39
MED23	2	1.4	41	26	33	21
MED25	0	0	10	10	8	7.4
MED27	0	0	8	27	5	23
MED31	0	0	6	32	7	34
CYCLINC	0	0	9	27	5	12
CDK8	0	0	5	4.5	2	4.5

Table 2: Isolation of Mediator.

Mediator components isolated with TAP-ICP4 at 3, 6, and 12 hpi. Numbers represent the percent coverage. Colors signify the Mediator module to which the component belongs (red= head, orange=middle, yellow=tail, green=kinase). Med1 is thought to act as a bridge between the middle and tail regions.

To verify the results obtained from LC-MS-MS, independent samples were subjected to western blot followed by protein quantification analyses. Figure 25 A represents the results of

the western blot analyses while Figure 25 B represents the results from the protein quantification analyses. For the purposes of directly comparing one protein between different samples, the spectral counts obtained from LC-MS-MS, or the number of individual peptides corresponding to the protein, are listed above the relative abundances of each protein (Figure 25 B). The western blot data confirmed that TAP-ICP4 was purified at both 3 hpi and 6 hpi as evidenced by the presence of immunoreactive bands in the TAP-ICP4 lanes, but not the KOS lanes (Figure 25A, panel 1). To account for the differences in quantity of ICP4 purified in each sample, protein quantification for all other proteins (TAF1, TBP, MED1) were normalized to the amount of ICP4 isolated within that specific sample. Western blot analyses indicated that both TAF1 and TBP were present in TAP-ICP4 purified samples at 3 hpi and 6 hpi, as determined by the presence of immunoreactive bands present at around 250 and 37kDa respectively (Figure 25 A panels 2 & 3). In accordance with LC-MS-MS data, protein quantification analyses indicated that TAF1 and TBP were more abundant in the 3 hpi sample than in the 6 hpi sample, with a 2-3 fold reduction in levels of TAF1 or TBP at 6 hpi compared to 3 hpi (Figure 25 B). Additionally, MED1 was not detectable by western blot analysis in TAP-ICP4 samples purified at 3 hpi, but was detectable at 6 hpi (Figure 25 A panel 4). Protein quantification analyses indicated a 3 fold increase in the amount of MED1 present in samples collected at 6hpi compared to 3 hpi, confirming the LC-MS-MS data. Together these data suggest that ICP4-mediated TFIID interactions preceded those of Mediator.

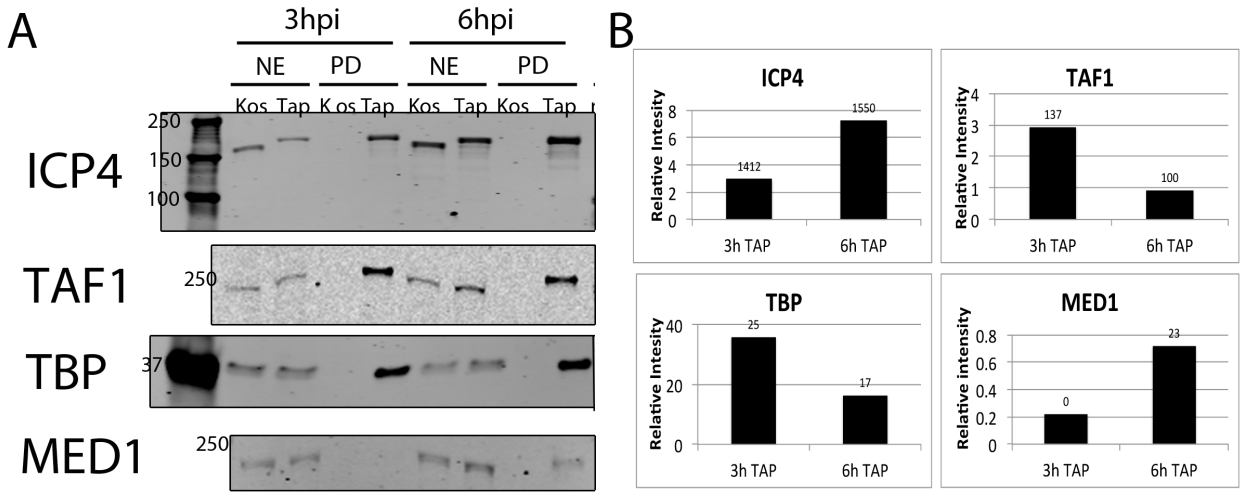


Figure 25: TFIID Quantification

A Western blot analyses of ICP4, TAF1, TBP, and MED1 from nuclear extracts (NE) and streptavidin affinity pull downs (PD). Samples were harvested from KOS and TAP-ICP4 (Tap) infected cells at 3hpi and 6hpi as labeled above. **B** Quantification of A, expressed as relative intensity.

The C-terminal Transactivation Domain of ICP4 Stabilizes Interactions with TFIID and Mediator

The viral mutant n208 lacks the C-terminal transactivation domain rendering it severely deficient for L gene expression (70). To determine whether n208 interacted with the same components of the transcription machinery as wtICP4 *in vivo*, TAP-n208 was purified at 6 hpi, the co-purified proteins were identified via LC-MS-MS, and compared to proteins isolated with TAP-ICP4 at 6 hpi.

The purification of components of TFIID and Mediator with TAP-ICP4 and TAP-n208 were compared (Tables 3 and 4). Notably, all 12 TAFs and TBP were isolated with TAP-n208, albeit some TAFs were less abundant compared to when isolated with TAP-ICP4. Comparison of the core components of TFIID indicated that the abundance of some TAFs in TAP-n208 samples was decreased, while the abundance of others remained approximately equal. For instance, TAFs 4 and 12 appeared to be present at similar levels in both the TAP-ICP4 and TAP-

n208, while TAFs 5 and 6 were decreased in TAP-n208 samples compared to TAP-ICP4 samples and TAF9 was significantly increased (Table 3). The decreased levels of some, but not all, core components of TFIID in TAP-n208 samples could be indicative of direct interactions between the N-terminus of ICP4 and certain core components of TFIID such as TAF9, however this remains to be experimentally determined. Of note, many TAFs were present in non-TFIID associated complexes suggesting that the differences in abundance of some TAFs may be the result of interactions, not only between ICP4 and TFIID, but also other TAF containing complexes. Interestingly, of the non-core TAFs, both TAF1 and TBP were less abundant in TAP-n208 samples compared to TAP-ICP4 samples as assessed by both the percent coverage and spectral counts. The reduced levels of TAF1 present in the TAP-n208 purified samples is in accordance with previously published data indicating that the C-terminus of ICP4 is necessary for TAF1 mediated interactions (41). It is perhaps not surprising that levels of TBP were also reduced in TAP-n208 samples, as TAF1 and TBP are known to directly interact (217, 317, 376); without ICP4-TAF1 interactions, stabilization of TFIID may not occur.

	TAP 6h		TAPn208 6hpi	
	SpC	% Coverage	SpC	% Coverage
TAF1	100	28	26	13
TAF2	67	34	91	40
TAF3	8	6.2	6	9.7
TAF4	43	50	37	56
TAF5	67	55	25	21
TAF6	76	55	11	14
TAF7	7	21	3	11
TAF8	11	22	12	39
TAF9	12	31	20	64
TAF10	8	25	7	18
TAF11	0	5.2	0	9.5
TAF12	14	27	14	21
TBP	17	28	4	3.7

Table 3: TFIID components affinity purified with n208.

TFIID components co-purified with TAP-ICP4 or TAP-n208 at 6 hpi. Numbers represent the percent coverage of the identified protein.

A comparison between the purification of components of the Mediator complex in TAP-ICP4 and TAP-n208 samples revealed that, as with TFIID, the C-terminus of ICP4 may not be necessary for interactions with Mediator, but instead may have a combined function with the N-terminus in stabilizing Mediator-ICP4 interactions. Many components of the Mediator complex were isolated at similar levels in TAP-n208 purified samples compared to TAP-ICP4 purified samples, including MEDs 4, 6, 14, 18, 20, 23, 27, and 31 (Table 4). Additionally, many components of the Mediator complex were reduced in TAP-n208 purified samples, including MEDs 1, 7, 12, 13, 15, 16, 17, 23, 25, 31, cyclin C, and cyclin dependent kinase (CDK) 8. Interestingly, many of the components that were reduced in abundance in the TAP-n208 sample (8/12) belong to the tail and kinase region of the Mediator complex, perhaps suggesting an interaction between the C-terminal transactivation domain of ICP4 and a component of the tail or kinase region of the mediator complex. This would be in accordance with published data indicating that co-activators frequently contact the tail region of the Mediator complex, allowing the head region to interact directly with RNA PolII.

	TAP 6h		TAPn208 6h	
	SpC	% Coverage	SpC	% Coverage
MED1	23	17	9	5.4
MED4	8	30	9	20
MED6	9	22	10	38
MED7	8	25	0	0
MED12	58	17	14	5.8
MED13	23	9.3	3	1.2
MED14	30	18	40	30
MED15	8	11	0	0
MED16	22	18	2	4.1
MED17	16	29	0	0
MED18	7	18	7	14
MED20	14	28	18	30
MED23	41	26	33	17
MED25	10	10	0	0
MED27	8	27	7	29
MED31	6	32	2	20
CYCLINC	9	27	0	0
CDK8	5	4.5	0	0

Table 4: Mediator components isolated from n208 purified samples.
Mediator components co-purified with TAP-ICP4 and TAP-n208 at 6hpi. Numbers represent the spectral counts (SpC) or the percent coverage (% Coverage) of the identified protein.

To verify results obtained from LC-MS-MS regarding TAP-n208 mediated interactions with TFIID and Mediator, western blot and protein quantification analyses were performed on independently isolated samples. Western blot analyses were performed with antibodies for ICP4, TAF1, TBP, and MED1. As evident by the presence of immunoreactive bands at approximately 250, 37, or 220 kDa in the TAP-n208 PD lanes, TAF1, TBP, and MED1, respectively, were co-precipitated in TAP-n208 samples (Figure 26 A). Protein quantification data indicated that 1.3 fold more n208 was purified during the procedure, and as such, all further protein quantification was normalized to the quantity of the ICP4 molecule purified. Protein quantification confirmed that TAF1, TBP, and MED1 were isolated to a lesser extent than with TAP-ICP4 (Figure 26 B).

Western blot and protein quantification analyses were in accordance with LC-MS-MS data and indicated that both TFIID and Mediator were isolated in TAP-n208 purified samples, but to a lesser extent than in TAP-ICP4 purified samples.

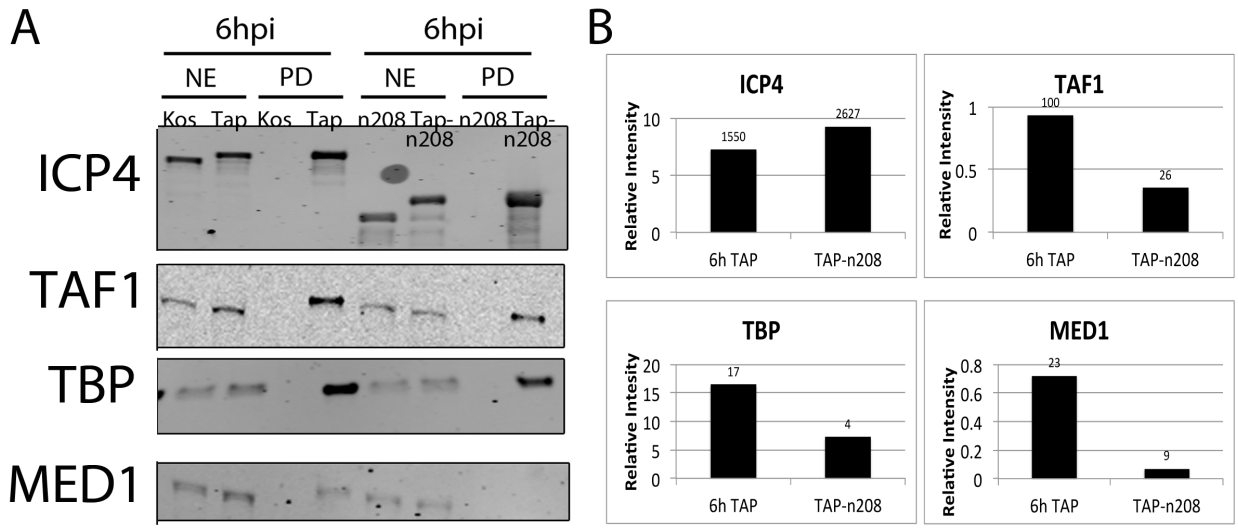


Figure 26: TFIID and Mediator purification from n208 infected cells.
A. Western blot analyses of KOS, TAP-ICP4, n208, and TAP-n208 purified samples. Antibodies directed against ICP4, TFIID (TAF1 and TBP), and Mediator (MED1) were utilized to determine the relative abundance of these protein complexes in purified samples. B. Quantification of A. Spectral counts as determined by LC-MS-MS are listed above the relative intensities.

Elongation Factors

Interestingly, in addition to cellular components involved in transcription initiation, many cellular components related to transcription elongation were co-purified with ICP4, including components of TFIIH and the SIII complex (Elongin). TFIIH is a multi-subunit protein involved in promoter clearance (107, 111, 211, 291). It is composed of a "core" including the XPB, p44, p62, p52, and XPD subunits and a dissociable "catalytic activating kinase" (CAK) region composed of MAT1, cyclinH, and CDK7. All 5 core components of TFIIH were isolated with purified ICP4, while cyclinH and CDK7 of the CAK were not (Figure 27 A). The kinetics of ICP4-TFIIH interactions were similar to those of the Mediator, with more components being

present and at higher levels later in infection (6 and 12 hpi) compared to early (3 hpi). As with TFIID and Mediator, components of TFIIH also co-purified with n208, although these components generally appeared to be less abundant in TAP-n208 purified samples compared to TAP-ICP4 purified samples. The presence of an immunoreactive band at approximately 62 kDa in a western blot analysis with antibodies directed at the p62 component of TFIIH verified that p62 was isolated with ICP4 and n208 at 6 hpi (Figure 27 B).

A

	TAP 3h	TAP 6h	TAP 12h	TAPn208 6h
XPB	0	8/23	16	5/5
p44	0	16	6	4
p62	0	2/12	7	2/12
p52	0	3/11	9	11
XPD	11	13	0	2/0
MAT1	0	5	0	0

B

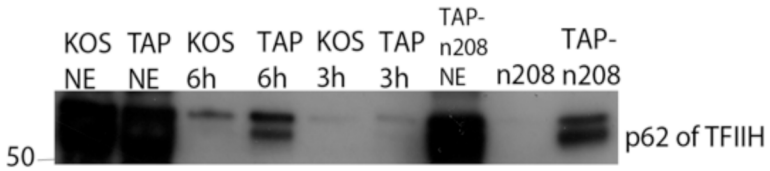


Figure 27: TFIIH is isolated in ICP4 containing complexes.
A. Components of TFIIH isolated during purification of ICP4, represented as the number of spectral counts. The number prior to the backslash represents the spectral counts in the untagged sample. No backslash indicates protein was not detected in the untagged sample. **B.** Western blot analysis using an antibody directed at the p62 component of TFIIH.

In addition to TFIIH components, the SIII (Elongin) complex also co-purified with TAP-ICP4 and TAP-n208. The SIII complex is involved in releasing or overcoming RNA PolII pausing on elongating RNA transcripts (318). The SIII complex is composed of three subunits: Elongins B and C, which are the regulatory subunits, and Elongin A, which is associated with transcriptionally active RNA PolII (26, 27). All three components of the SIII complex co-purified with TAP-ICP4 late in infection (6 and 12 hpi), while only the regulatory subunits,

Elongins B and C, co-purified with TAP-n208 (Table 5). While Elongins A and C were also present in the untagged samples at 3 and 6 hpi, Elongin A was enriched more than 5 fold in the TAP-ICP4 sample at 6 hpi.

	KOS 3h	KOS 6h	KOS 12h	n208 6h
	TAP 3h	TAP 6h	TAP 12h	TAP-n208 6h
ElonginA	10/0	3/16	13	0
ElonginB	0	3	0	5
ElonginC	6/8	4/2	12	11

Table 5: Isolation of the Elongin Complex.

Components of Elongin SIII isolated with purified ICP4, represented as number of spectral counts. The number prior to the backslash represents spectral counts in the untagged sample, whereas the number after the backslash represents the spectral counts in the TAP-tagged sample. No backslash indicates that the protein was not identified in the untagged sample.

In an attempt to identify regions of ICP4 that are responsible for direct protein-protein interactions, a small highly conserved, biologically relevant 20 amino acid region corresponding to amino acids 80-100 within the N-terminus of ICP4 (340) was engineered to contain a GST tag (m90GST). The 20 amino acid peptide was then bound to GST-conjugated beads and incubated with HeLa nuclear extract. Non-specific proteins were washed away and the GST-m90 peptide with bound proteins was eluted from the beads using 10 mM glutathione. Eluted samples were concentrated and analyzed via LC-MS-MS and western blot. Interestingly, the bound m90 peptide was capable of interacting with Elongins B and C (Figure 28). Elongin A was not found in either the GST-purified m90 samples or in n208 containing complexes (Table 5 and Figure 28). These data suggest that perhaps the C-terminus of ICP4 is important for interactions with Elongin A. Elongin A is the transcriptionally active component of the SIII complex, therefore, the reduction in transcription of E and L genes displayed by n208 could potentially be due to an inability of the molecule to recruit/stabilize Elongin A. This data, combined with the TAP-ICP4

purification data (Table 5), suggest that the SIII complex is a *bona fide* component of ICP4 containing complexes, although further experimentation is needed to verify this.

A

	GST	m90-GST
Elongin B	0	10
Elongin C	0	14

B

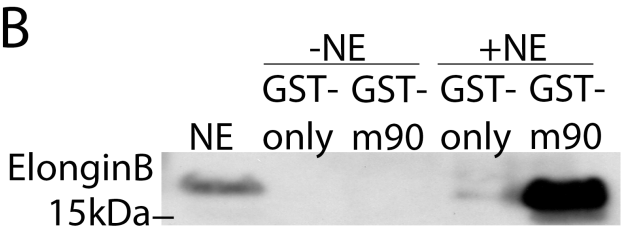


Figure 28: Elongin SIII purification with m90-GST.
Elongin components isolated with GST by itself or m90-GST, represented as spectral counts. Samples containing GST only or GST-m90 were bound to GST conjucated beads, and either incubated with HeLa nuclear extract (+NE) or not (-NE). Samples incubated with nuclear extract were analyzed via LC-MS-MS. A. LC-MS-MS results, represented in spectral counts. B. Western blot analysis with antibodies directed against Elongin B.

Chromatin Remodeling Factors:

In addition to the many components related to transcription initiation and elongation, multiple components of the SWI/SNF, Nurd, and Ino80 chromatin remodeling complexes were isolated in ICP4 containing complexes. Chromatin remodeling complexes hydrolyze ATP to provide the energy to disrupt histone/DNA interactions, thus allowing them to move, slide, or replace histones on the DNA (Reviewed in 47, 126). As such, a common characteristic of chromatin remodeling complexes is the presence of an ATPase domain. The ATPase containing proteins of the SWI/SNF, Nurd, and Ino80 complexes are Brahma related gene 1 (Brg1) & Brahma (Brm), CHD3&4, and RUVBL1&2, respectively. Interestingly, with the exception of Brg1, each of these ATPases was co-purified with ICP4 from infected cells (Figure 29 A, C, D). The co-purification of Brahma was verified using western blot analysis (Figure 29 B). The

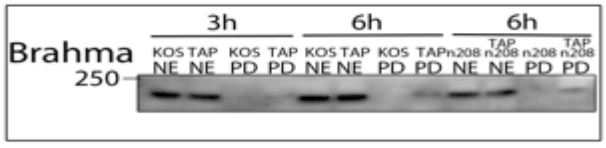
presence of an approximately 220 kDa immunoreactive band in the TAP-ICP4 6 hpi and TAP-n208 samples validated the LC-MS-MS identification Brahma. In addition to the ATPase containing proteins, multiple adaptor proteins from each of these complexes were isolated. These included SMARCA4, SMARCC2, and Bzb1 kinase from the SWI/SNF complex, HDAC2 from the Nurd complex, and Apr5 from the Ino80 complex (Figure 29).

With respect to the kinetics of ICP4-chromatin remodeling complex interactions, the Ino80 complex appeared to be more prevalent late in infection (Figure 29 D). This is perhaps not surprising, as the Ino80 complex has been implicated in stabilizing replication forks during DNA replication (249, 297). The recruitment of Ino80 to ICP4 containing complexes may be the result of the recognition of branched concatemers as replication forks. This is supported by the fact that Ino80 components are found to higher extents later in infection (6 and 12 hpi) when viral DNA is replicating, and not in cases where DNA replication does not occur such as early in infection or in n208 infected cells (Figure 29 D).

A

Swi/Snf				
	TAP 3h	TAP 6h	TAP 12h	TAP-n208
SMARCA4	3	9	4	0
SMARCC2	0	0	5	6
Brahma	0	8	0	4
BZB1 kinase	2/16	0	14	5

B



C

Nurd				
	TAP 3h	TAP 6h	TAP 12h	TAP-n208
CHD3	18	12	34	7
CHD4	13	0	0	4
HDAC 2	7	4/0	0	0

D

Ino80				
	TAP 3h	TAP 6h	TAP 12h	TAP-n208
RuvBL1	7/6	23	22	8/12
RuvBL2	6	2/25	25	5/11
Arp5	0	5	10	3

Figure 29: Chromatin remodelers isolated in ICP4 containing complexes.

A. Components of the SWI/SNF complex. B. Western blot analysis using antibodies directed against Brahma of the SWI/SNF complex. C. Components of the Nurd complex identified via LC-MS-MS. D. Components of the Ino80 complex. Data is represented as the number of spectral counts identified for each protein. The numbers listed prior to the backslash represent the number of spectral counts in the untagged samples.

ICP4 and mRNA processing: Capping to Export

In addition to components involved in chromatin remodeling, transcription initiation, and elongation, many components related to RNA processing were also isolated in ICP4 containing complexes (Table 6). Twenty-six protein components constituting a portion of the spliceosome were isolated in ICP4 containing complexes. This was surprising, as the vast majority of HSV-1 genes are not spliced (339). The isolation of the spliceosome components with ICP4 may reflect either the isolation of ICP4 containing complexes involved in cellular gene regulation, or conversely may suggest that ICP4 sequesters components of the splicing machinery to aide in the inhibition of cellular splicing. It has been shown that in HSV-1 infected cells, a Ser2 phosphorylated version of RNA PolII localizes to nuclear foci containing components of the spliceosome (108). It was also shown that some of these sites are juxtaposed to ICP4 containing foci (108), providing a potential link between ICP4 and splicing factors.

Additionally, components involved in the cleavage of nascent RNA transcripts were isolated. These include cleavage and polyadenylation specificity factors (CPSF) 1, 2, 3L, 4, 7 and FIP1L1. Multiple components involved in mRNA export were isolated, including members of the TREX complex. The overwhelming number of components related to mRNA processing isolated in ICP4 containing complexes, in part, substantiates itself, however, western blot analyses will provide further validation of this data. The presence of the numerous components involved in gene transcription, but the absence of certain notable proteins such as RNA Polymerase II, TFII A, B, E, and F, other strong DNA binding proteins, and viral proteins leads

us to believe that the majority of the proteins identified are *bona fide* components of ICP4-mediated transcription complexes. The potential roles of the mRNA processing factors, chromatin remodeling complexes, and elongation factors in ICP4 mediated viral gene expression will be discussed at length.

	Function	**TAP 3**	**TAP 6**	**TAP 12**	**TAP-n208**
Poly-Adenylation					
CPSF4	makes mature mRNA; cleavage	2/9	0/8	0/7	0/16
CPSF2	makes mature mRNA; cleavage	6/16	0/12	3/31	9/0
CPSF1	makes mature mRNA; cleavage	9/20	5/18	5/31	22/51
CPSF7	makes mature mRNA; cleavage	4/6	0/8	0/6	7/10
CPSF3 like	makes mature mRNA; cleavage	0/11	4/8	2/10	6/3
FIP1L1	subunit of CPSF; stimulates poly(A)polymerase activity	0/8	0/9	0/20	9/7
Splicing					
HCC-1	Coactivator involved in splicing	11/9	4/4	0/23	4/4
Pre-mRNA processing factor 6	bridge between U5 and U4/U6	3/6	6/36	4/46	4/7
Pre-mRNA processing factor 38	spliceosome assembly	6/9	0/5	0/7	0/13
splicing factor 1	spliceosome assembly	9/20	5/18	5/18	4/6
snRNP40	component of U5 spliceosome	0/6	0/14	0/23	0/9
PRPF4	component of U4/U6 spliceosome	0/0	0/9	0/3	6/6
hnRNP A/B	many functions; part of the spliceosome; binds to tk promoter	4/9	6/11	0/7	0/8
SmD1 like protein	spliceosome	9/21	6/9	3/27	6/8
Luc7-like (similar to U1)	spliceosome	3/6	0/0	0/12	0/20
ETFUD2	component of U5	33/19	31/45	15/78	27/65
U4/U6.U5 tri-snRNP associated protein	involved in U2 spliceosome splice site recognition	0/0	0/14	0/41	4
RNA binding protein 10	regulation of alternative splicing	9/4	8/3	0/5	0/0

pre-mRNA banch site protein p14	contacts branch site for first catalytic step of splicing	8/8	0/0	0/5	0/5
RNA Binding protein 25	alternative splicing/site selection	4/18	4/62	4/24	4/107
pre-mRNA processing factor 8	spliceosome; aligns 5' and 3' exons for ligation	33/87	33/139	20/212	56/100
Serine/arginine rich splicing factor 4	alternative splice site selection	17/11	13/0	0/10	0/0
snRNP D2	spliceosome	7/7	0/2	0/8	0/15
snRNP200	component of U5	37/86	38/128	15/165	54/77
PUF60	splice site recognition; also inhibits active txn via TFIIH	8/38	8/92	0/119	8/148
U2AF	component of U2; binds to pre-mRNA branch site	22/51	14/57	0/51	2/174
Imp3	component of U3; spliceosome	12/5	4/0	0/10	4/4
splicing regulatory glutamine/lysine rich protein1	regulates other splicing factors	0/3	0/10	0/13	0/21
snRNP E	part of U7 complex; 3' histone processing	3/4	0/4	0/7	0/3
PRPF3	associates with U4/U6	0/0	0/10	0/7	2/3
splicing factor 3B subunit2	anchors U2 to the branch site	10/6	17/4	0/13	4/6
snRNPB	component of U2	11/5	2/4	0/8	2/15
Export					
T-rex complex 1	packages mRNA for export	0/7	0/7	0/3	0/4
Mahog	component of EJC	10/8	3/0	0/14	0/0
hnRNP L-like	mRNA packaging/export	16/8	12/6	0/6	12/6
Nuclear Cap binding protein subunit 2	component of CBC; interacts with ALYREF/T-rex; meditates intronless mRNA export	7/5	5/6	0/6	4/2
T-rex complex 6	packages mRNA for export	0/7	2/9	0/0	0/0
nuclear cap binding protein subunit 1	component of CBC; interacts with ALYREF/T-rex; meditates intronless mRNA export	11/11	3/3	0/10	11/0
RNA Helicases					
DDX50	rRNA synthesis/processing	13/12	0/0	0/11	10/0
DDX6	P-body associated; translation	5/4	0/5	0/4	6/0

	suppression and mRNA degradation				
DDX23	unknown	0/0	0/0	3/0	0/0
DDX47	shuttles between nucleus and cytoplasm	24/20	6/15	5/9	20/9
DDX41	intracellular DNA sensor	27/50	9/53	5/50	17/62
DDX39B	splicing factor required for U2 association with pre-mRNA	8/5	9/9	3/10	11/0
DDX23 like	component of U5 snRNP	8/8	3/24	3/44	11/23
DDX49	unknown	4/3	0/5	0/2	0/3
DDX8	releases RNA from spliceosome to facilitate nuclear export	0/0	0/0	0/8	4/37
mRNA stability					
Serpine1 mRNA binding protein	binds 3'end of RNA to stabilize mRNA; interacts with CHD3	0/3	0/6	0/5	6/9
Mahog	component of EJC	10/8	3/0	0/14	0/0
Unknown					
RNA binding protein 26	unknown	4/0	0/0	0/9	0/8
RNA binding protein 15	unknown	19/8	8/9	0/7	14/2
Nop5/Nop58	methylation of rRNA?	36/35	25/21	8/43	47/30
pre-mRNA processing protein 40A	splicing?	8/19	2/66	4/17	3/127
Luc7 Like2	RNA binding	16/42	4/60	0/73	0/197
Serine/Arginine rich splicing factor 11	may play a role in pre-mRNA processing based on structure	0/11	0/11	0/18	0/33

Table 6: mRNA processing factors isolated in ICP4 containing complexes.

Data is represented as spectral counts. Numbers prior to backslash are from untagged samples, numbers post backslash are from TAP-tagged samples.

4.5 DISCUSSION

The purposes of these studies were to i) determine the temporal kinetics of ICP4 mediated interactions, ii) determine the structural requirements for these interactions, and iii) identify novel proteins in ICP4 containing complexes. Affinity purification of ICP4 was performed, and co-purified proteins were identified via LC-MS-MS and verified via western blot analyses. Here, we have shown that ICP4 mediated interactions with TFIID precede those of Mediator, and that the N-terminal 774 aa of ICP4 are sufficient, but sub-optimal, for both of these interactions. Additionally, we show that ICP4 containing complexes also contain proteins involved in chromatin remodeling, elongation, and mRNA processing and export.

Time Couse Analyses of Transcription Complex Assembly

The data presented here strongly suggest that ICP4 interacts with TFIID prior to the Mediator complex. This may be reflective of the kinetics of PreInitiation Complex (PIC) formation on E viral promoters. During transcription initiation, TFIID, facilitated by TFIIA, binds to a core promoter element, such as a TATA box or Inr element. This triggers the subsequent assembly of TFIIB, PolII and TFIIF, and TFIIE on the promoter to form a functional PIC (Reviewed in 275). It was originally demonstrated that TFIID binding to a core promoter element is a rate-limiting step in PIC assembly (357). More recently, the Roeder lab demonstrated that the interaction of TFIID and Mediator on the promoter is the rate-limiting step for the initiation of basal transcription, and suggested that a TFIID-Mediator complex must be present on promoter DNA to initiate basal transcription (7). Studies investigating activator dependent recruitment of Mediator, indicate that Mediator is not always recruited to activated genes (96), suggesting, that TFIID-Mediator interactions may not always be the rate limiting step in the initiation of activator dependent transcription.

Our lab has shown that TFIID and Mediator are present on viral promoters, and that their recruitment is dependent on the presence of ICP4, thus demonstrating that viral transcription occurs in an activator (ICP4)-Mediator dependent manner (198, 284). Data establishing that ICP4 and TFIID form complexes prior to interactions with the Mediator complex suggest that ICP4-TFIID complexes are likely formed early, presumably on DNA, prior to other components of the initiation machinery. The mechanism of Mediator recruitment to viral promoters has not been elucidated. It is possible that ICP4-TFIID interactions may be required for the recruitment of Mediator to viral transcription complexes. ICP4 likely plays an integral role in the recruitment of the Mediator complex, as other co-activators were not isolated in ICP4 containing complexes, despite the presence of several activator-binding sites within the promoter region of E genes. Nonetheless, the possibility that other cellular/viral factors or post-translational modifications, promote the recruitment of Mediator to viral genomes cannot be ruled out.

Chromatin Remodeling, and Late Gene Expression.

The ICP4 mutant n208 is missing the C-terminal 500 amino acids corresponding to the C-terminal transactivation domain, yet retains the DNA binding domain, NLS, dimerization motif, and N-terminal transactivation domain. n208 is capable of activating E, but not L gene expression, and does not replicate DNA efficiently. The data herein demonstrated that n208 is capable of interacting with TFIID and Mediator, yet these interactions are not sufficient for DNA replication or L gene expression. This is perhaps reflective of a presumably lower affinity interaction between n208 and TFIID and potentially Mediator. The C-terminus of ICP4 has been established as being important for interactions with TAF1 of TFIID (41). Consequently, in n208-infected cells approximately 4 fold less TAF1 was isolated. Additionally, TAF1 and TBP have been shown to directly interact (217, 317, 376), thus, approximately 4 fold less TBP was

also isolated from n208 infected cells. The reduced level of isolation of TFIID in n208 containing complexes may potentially result in unstable interactions with the Mediator complex, as we and others have suggested that TFIID-Mediator interactions may be a limiting step in the initiation of transcription (7). The ability of n208 to interact with TFIID at all is presumably a result of the redundant functions encoded by ICP4; the N-terminus of ICP4 also stabilizes interactions with TFIID through TBP (303, 340, 373).

The inability of n208 to activate L gene expression may be due to reduced stabilization of TFIID and mediator on viral promoters, however, as viral DNA replication is dependent on E gene transcription, in addition to ICP4-TFIID interactions, other factors and mechanisms may be involved. One such factor may be the Ino80 complex. In this system three components of the Ino80 complex, RuvBL1, RuvBL2, and Arp5 were isolated in ICP4 containing complexes, but not in n208 containing complexes. Additionally, significant amounts of the Ino80 complex were not isolated until 6 hpi, after the onset of vDNA replication in this system. The Ino80 complex contributes to DNA replication fork progression, and stabilization (249), presumably through remodeling chromatin components associated with the replicating DNA (Reviewed in 56). While this is the first study implicating the Ino80 complex in HSV DNA replication, the mechanisms underlying chromatin remodeling, transcription, and DNA replication are still being elucidated. Recent advances in chromatin dynamics during lytic infection indicate that viral DNA associates with histones, histone variants, and SWI/SNF chromatin remodeling complexes (Reviewed in 212). In addition to its role in DNA replication, the Ino80 complex has also been shown to promote the activation of transcription via chromatin remodeling (56). In this manner, the Ino80 complex may have a dual function in both transcription and DNA replication during HSV lytic infection. Based on these findings we propose a model in which the C-terminus of ICP4 may be

involved in promoting viral DNA transcription and replication, perhaps through its ability to multimerize on DNA (183) and interact with components involved in chromatin remodeling and replication fork stability.

Gene Looping.

Perhaps one of the most interesting findings in this study is the association of ICP4 with components involved in mRNA processing and export. The shear number of proteins isolated that are involved in mRNA processing indicates that ICP4, which was primarily viewed as a DNA dependent transcription factor, may play an integral role in mRNA processing. Within the last decade the view of transcription and mRNA processing has evolved to support a model in which the steps of mRNA processing are intimately linked co-transcriptionally (226). Many actively transcribing genes undergo a process referred to as "gene looping". After a productive round of transcription, RNA PolII is released from the terminator and can re-associate with a reinitiation scaffold at the promoter that is comprised of several GTFs and activators. By linking the terminator and promoter regions of genes, hence forming a gene loop, reinitiation is made more efficient (Reviewed in 292). The cleavage and polyadenylation specificity factors (CPSFs), which are involved in cleaving the polyadenylated nascent transcript, have been intimately linked to gene looping as they provide an interface for interaction with several GTFs and coactivators including TFIIB, TFIIH, and PC4 (36, 221). PC4 is one of very few proteins that contain sequence similarity to ICP4 (110). PC4 interacts with CPSF and TFIIB, thus providing a direct link between the promoter and terminator regions of actively transcribing genes (36). Additionally, mRNA processing and gene looping have been shown to be coupled to mRNA export. Based on the similarity between ICP4 and PC4 and the purification of multiple CPSFs and mRNA export proteins, we propose a new model for ICP4 mediated transcription in

which ICP4 directly interacts with CPSFs providing a structural link between the promoter and terminator region of HSV genes (Figure 30). The function of gene looping in HSV transcription may be to increase the efficiency of re-initiation, although HSV genes are not particularly long as they are generally intronless. As such, the formation of gene loops may not be necessary for HSV transcription, and may be a consequence of the transcription process. Interestingly, integrated HIV provirus also forms gene loops, which enhance transcription rates (256). During HIV transcription, elongation appears to be an integral part of the formation of gene loops, as initiation is not sufficient to drive their formation. The data herein simply suggest that ICP4 may bind to CPSFs to establish a gene loop. However, direct evidence demonstrating gene looping and its function in HSV infection needs to be established, and are currently under investigation.

The data presented herein suggest that ICP4 is involved in a variety of processes related to transcription. Given the large size and elongated structure of the molecule, ICP4 likely provides a binding surface for a variety of transcription complexes. In addition, many of the aforementioned cellular transcription processes are coupled to each other, thus increasing the rates of cellular transcription. Activated transcription from the viral genome requires a delicate balance between the mechanisms evolved by the host cell to combat viral infection and the molecular mechanisms developed by the virus to overcome these barriers. It is not surprising, given the size and structure of ICP4, that ICP4 would be intimately linked to many of these processes.

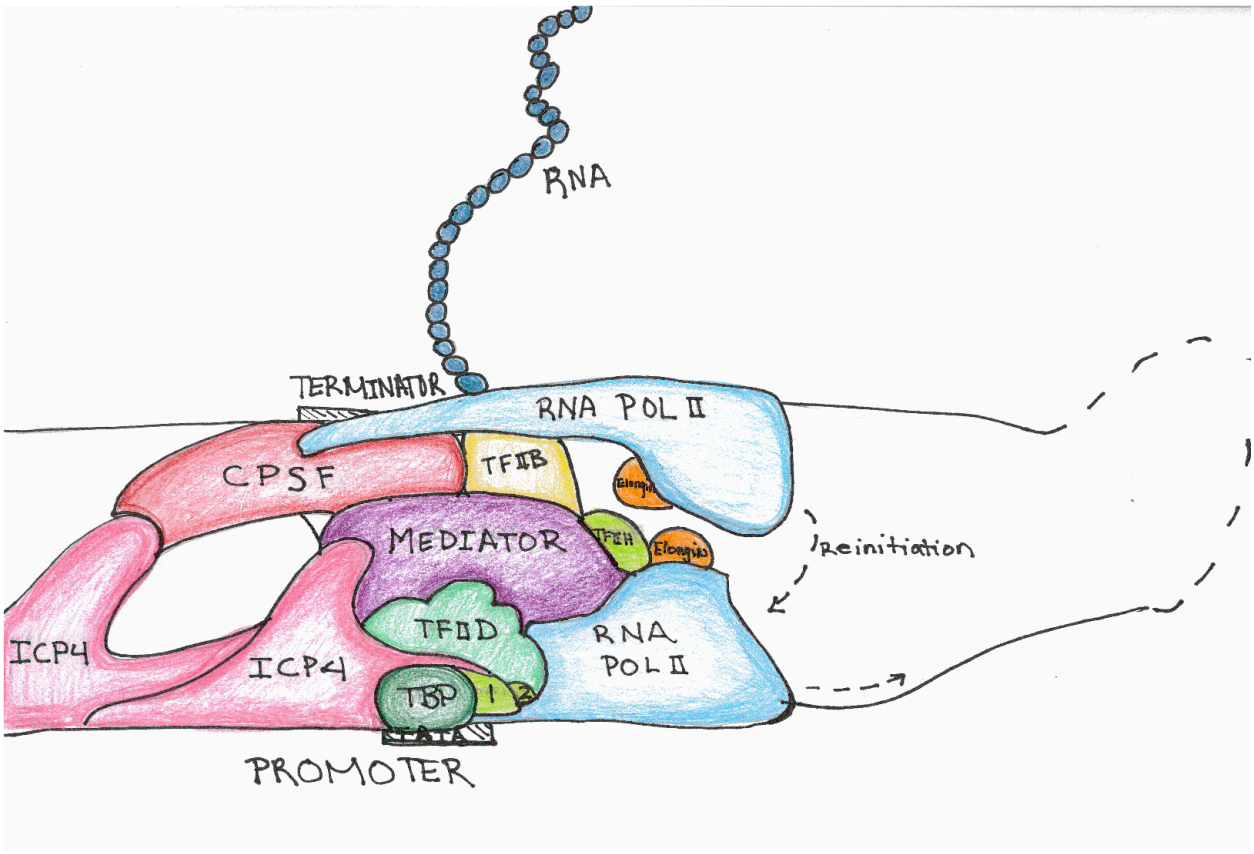


Figure 30: Model of ICP4 mediated gene looping.
ICP4 interacts with CPSF forming a structural link between the promoter and terminator regions of HSV genes. This process may influence reinitiation kinetics.

4.6 ACKNOWLEDGMENTS

The construction of the m90-GST tagged construct and the associated purifications and western blot analyses were performed by the skilled technician, Fran L. Sivrich.

5.0 THESIS SUMMARY AND GENERAL DISCUSSION

5.1 SUMMARY OF THESIS

ICP4 functions at the level of transcription and is capable of interacting with components of the transcription machinery, including TFIID and Mediator (198). Two transactivation domains and a DNA binding domain help mediate the transcriptional activities of ICP4. While the contributions of the transactivation domains to viral transcription have been investigated, a comprehensive understanding of how these domains function individually and together has not been established. We hypothesized that regions within both transactivation domains contribute individually and together to the activities of ICP4. Therefore, the goal of this project was to more accurately define the contributions of the transactivation domains to the activities specified by ICP4. Given that ICP4 functions through transcriptional complexes, the complexity of ICP4 mediated interactions throughout infection was also investigated.

To investigate the contribution of the transactivation domains to ICP4 activities, deletion mutants spanning the N-terminal transactivation domain were constructed in the presence and absence of the C-terminal transactivation domain. The contributions of these regions were assessed with respect to viral growth and the kinetics and levels of mRNA and polypeptide expression. The data demonstrated that in the absence of the N-terminal transactivation domain, viral transcription was abolished and ICP4 was no longer found in complexes with TFIID.

Additionally, the data suggested that interactions with TFIID could be restored by the formation of a heterodimer consisting of a N-terminal deletion mutant and a C-terminal deletion mutant, ultimately resulting in restored viral growth. Further characterization of the N-terminal transactivation domain failed to yield specific regions that were necessary for viral growth, although some regions specified functions related to the levels or kinetics of transcription in cell culture. Multiple independent regions affected the levels of viral growth and gene expression to varying degrees, indicating that interactions with transcriptional machinery likely occur on multiple surfaces of the molecule. The deletion of amino acids 81-96 within the N-terminal transactivation domain had no effect on viral growth in a cell culture model, however, the additional deletion of the C-terminal transactivation domain resulted in attenuated viral transcription and growth, indicating that regions in the N-terminus and C-terminus both contribute to a specific function of the molecule.

ICP4 functions through transcription complexes. Therefore, one of the goals of this work was to more accurately define the transcription complexes with which ICP4 interacts and determine how these interactions change throughout infection. Affinity purification followed by LC-MS-MS and western blot analyses determined that ICP4 is involved in a variety of transcription complexes. The kinetic analyses indicated that ICP4 forms a complex with TFIID prior to other components of the transcription machinery such as Mediator or TFIIH. Additionally, proteins involved in chromatin remodeling, initiation, elongation, cleavage and polyadenylation, and nuclear export were isolated in ICP4 containing complexes indicating that ICP4 functions at multiple levels beyond transcription initiation.

5.2 GENERAL DISCUSSION

Structural Requirements for ICP4 Activities:

The work presented in this study provides insight into structural requirements for ICP4 mediated functions. Electromobility shift assays, predictions of disorder, and the values obtained for stokes radii suggested that ICP4 has an elongated structure with the N-terminus being more elongated than the C-terminus (293, 340). The elongated N-terminal transactivation domain of ICP4 is required for viral transcription and interactions with at least TFIID. Affinity purification and LC-MS-MS analyses indicate that n208 was capable of interacting with the same transcription complexes as wtICP4, although these complexes appeared less stable. Together this data suggested that the N-terminal transactivation domain may contribute significantly to the majority of ICP4 mediated protein interactions. The elongated structure of the N-terminus likely provides ample surface area for transcription complex binding and activators to function at a distance while the C-terminal activation domain is integral to stabilizing interactions between ICP4 and the transcriptional machinery, and perhaps in stabilizing the disordered N-terminal portion of the molecule upon factor binding, which will be discussed below.

Data indicated that the N-terminal transactivation domain of ICP4 was necessary for the associated transcriptional activities. Interestingly, an amino acid sequence comparison of ICP4 from multiple alphaherpesviruses demonstrates that there are two highly conserved regions corresponding to the DNA binding domain and the C-terminal transactivation domain (66). Thus, it is noteworthy that the conserved C-terminal transactivation domain is dispensable for E gene transcription while the more degenerate N-terminal transactivation domain is not. This is likely the result of the structure, or lack there of, of the N-terminus; it contains small highly conserved ordered regions that are flanked by largely disordered regions. It is likely that many

of these highly conserved regions provide binding surfaces for a variety of transcription complexes. The presence of disordered regions of ICP4 suggests that these regions are flexible, potentially providing for a variety of conformational changes upon factor binding. Affinity purification data suggested that ICP4 interacts with multiple protein complexes; the wide array of proteins with which ICP4 interacts could be partially dependent on changes in conformation of the flexible regions within the N-terminus. Additionally, the shear length of the molecule, provided largely by the N-terminus, may be important for allowing ICP4 to act at a distance, thus the transcription defects observed with deletion of amino acids 30-142 and 142-210 may be the result of not only the deletion of conserved binding sites, but also disruption of the length of the molecule. This could prevent ICP4 from interacting with activators at a distance, even if the regions responsible for those protein-protein interactions are retained.

It is particularly interesting that the N-terminal 774 amino acids of ICP4 are sufficient to interact with the same transcription complexes as wtICP4 and yet viral DNA replication and L gene expression are greatly diminished. Thus, the C-terminus must have a role in these functions. The reduced levels of viral DNA replication in n208 infected cells is particularly interesting because ICP4 has never been directly implicated in viral DNA replication. Presumably, the lack of DNA replication observed in n208-infected cells is the result of the decreased abundance of the E proteins involved in DNA replication. However, at the mRNA level, n208 accumulated similar levels of tk mRNA as either d3-8 or d8-10, which both replicate DNA and were only moderately reduced in viral growth. Comparisons between the levels of viral DNA replication and the quantities of E proteins produced, particularly those involved in DNA replication, would be of interest.

Another explanation for reduced levels of DNA replication in n208 infected cells arose from the affinity purification and LC-MS-MS data that indicated that n208 does not strongly associate with the Ino80 chromatin remodeling complex, while wtICP4 does. This complex has been linked to chromatin remodeling during cellular DNA replication in addition to replication fork stabilization (56). The association of ICP4 with the Ino80 complex may be the result of viral DNA replication, or conversely, viral DNA replication could require the association of the Ino80 complex. It would be intriguing if the C-terminus of ICP4 directed this complex to the viral DNA to allow for efficient DNA replication. The contributions of the Ino80 complex to HSV transcription and DNA replication have not yet been investigated. In this respect, it would be beneficial to determine i.) whether the Ino80 complex associates with the viral genome, ii.) if it localizes to either OriL or OriS during viral DNA replication, iii.) if it is necessary for viral DNA replication or transcription, and iv.) the impact of ICP4 on its association with the viral genome.

Interestingly, DNA replication is not a prerequisite for wtICP4 to bind to L promoters, yet n208 is not present at the gC promoter (284). The defects in L gene transcription are partially the result of the inefficient formation of transcription complexes on L promoters (284). The comparatively lower abundances of proteins isolated with n208 indicates a reduced affinity for these transcriptional complexes. Additionally, n208 does not multimerize on viral DNA, which results in a decreased affinity of ICP4 for DNA (183). The architecture of L promoters differs significantly from that of E promoters in that they are generally devoid of binding sites for cellular activators. Thus, the defects observed in L gene expression may arise from the decreased affinity of n208 for both transcription complexes and viral DNA, in addition to the lack of upstream activators to help stabilize ICP4 and transcription machinery on DNA.

As suggested above, the N- and C-terminal transactivation domains likely function together to form multiple protein contacts. Deletion of amino acids 81-96 (m90), which corresponds to a highly conserved, ordered region within the N-terminus had minimal effect on viral growth in cell culture. The additional deletion of the C-terminus of ICP4 (m90n7) rendered the molecule completely defective for viral transcription. This data implies that the N-terminus and C-terminus of ICP4 contain redundant functions. It is likely that these regions both act to stabilize a particular transcription complex, perhaps TFIID or Mediator, on viral promoters. It is probable that ICP4 contacts multiple subunits of both TFIID and Mediator. For example, the N-terminal transactivation domain is known to contact TBP while the C-terminal transactivation domain contacts TAF1. The presence of either of these interactions may be sufficient for the ICP4-TFIID interactions; however, the loss of both regions could ablate this interaction. Similar effects were observed with deletion of amino acids 30-142 (d3-8 and nd3-8) and 142-210 (d8-10 and nd8-10). Studies aimed at determining the functional deficiencies in m90n7, nd3-8 and nd8-10, compared to their full length counterparts would help delineate the surfaces of ICP4 that interact with the transcriptional machinery. It is striking that the virus has evolved to provide functional redundancies within the molecule in addition to maintaining two genomic copies of ICP4. This underscores the importance of ICP4 to the viral lifecycle.

When evaluating the structure of ICP4, the fact that ICP4 exists and functions as a dimer must be taken into account (223). It is established that coinfection with viral mutants differing in defective functions can lead to complementation between the two peptides (294). It has been shown that coinfection of n208 with a mutant that does not bind to DNA (d2) leads to a novel protein-DNA complex and that transcription of L genes is restored. Thus, the DNA binding defect in d2 was complemented by n208 and the functional C-terminal domain of d2

complemented the defects in L gene transcription exhibited by n208. The presence of novel protein-DNA complexes in coinfected cell extracts suggests that a heterodimer between n208 and d2 was present on the DNA (294). Additionally, it has also been established that the structure of one molecule within a heterodimer can negatively alter the conformation of the opposite molecule (296), suggesting that the structures of C-terminal and N-terminal transactivation domains directly affect each other. This implies that the N-terminal and C-terminal transactivation domains may be in close proximity to each other, and that the dimer may be structured in a head to tail orientation.

Data collected in these studies were also aimed at enhancing the understanding of the structure of the ICP4 dimer. An ICP4 mutant lacking the N-terminal transactivation domain, d3-10, failed to interact with TFIID resulting in undetectable levels of E and L gene transcription. Infection of d3-10 expressing cells with n208 resulted in complementation, with increased viral progeny, L gene expression, and the formation of d3-10-n208 heterodimers that were capable of interacting with TFIID. Studies herein demonstrated that n208 itself is capable of interacting with TFIID. Thus, the n208 molecule within the heterodimer may provide for the recruitment of TFIID while the C-terminus of d3-10 may allow for mulitmerization or recruitment of other complexes (perhaps Ino80) important in viral DNA replication and L gene transcription. Additionally, the C-terminus of d3-10 may act to stabilize the interactions with TFIID that are moderately unstable with n208 alone. This data implies that the N-terminus and C-terminus of opposite molecules within the dimer can function together to stabilize TFIID binding on L promoters. This reinforces the hypothesis that ICP4 dimers may be structured with the N-terminus and C-terminus of the opposite molecules close together.

ICP4: A hub for transcription complexes.

It has been established that ICP4 is involved in the formation of TFIID and Mediator complexes on viral promoters (198). Data presented herein suggest that ICP4 may function at more levels than transcription initiation. Affinity purification, LC-MS-MS, and Western blot analyses suggest that ICP4 interacts with multiple components of the transcription machinery, including those involved in initiation, elongation, mRNA processing, and mRNA export.

TFIID nucleating the core promoter is regarded as the initial step of PIC formation on TATA driven promoters, and as such, is required for most components of the transcription machinery to properly position on viral promoters. Interestingly, most of the ICP4 mutants studied to date either eliminate or reduce the stability of TFIID interactions, thus affecting complex formation with other factors, such as Mediator. A mutant that retains stabilized binding with TFIID, but not other transcription complexes has not been isolated. However, some deletion mutants display differing phenotypes with respect to both the levels and kinetics of E gene expression. In particular, the mutant d8-10 was capable of activating E gene transcription, but increased levels of mRNA transcripts did not accumulate until 8hpi. Conversely, the mutants d3-8 and n208 showed generally reduced levels of E gene transcription. n208 does not form particularly stable interactions with TFIID, therefore it can be hypothesized that the regions deleted in d3-8 are also involved in stabilization of TFIID. However, the differing phenotype of d8-10 suggests that amino acids 142-210 may have a distinct role in the formation of transcription complexes. In this respect, affinity purification of d3-8 or d8-10 and analyses of the associated transcription complexes would provide insight into mechanisms of ICP4 mediated transcriptional activation.

Upon the formation of transcription complexes on viral promoters, RNA PolII mediated elongation occurs. The CTD of RNA PolII is alternatively phosphorylated in HSV infection and

lacks the Ser2 phosphorylation (108) that is generally associated with the recruitment of mRNA processing factors. Alternative phosphorylation of RNA PolII appears to be mediated by the virus as a mechanism to subvert RNA PolII from cellular transcription processes to viral transcription processes (208). Interestingly, Med26, a variable component of the mediator complex, was not isolated with wtICP4 in infected cells. Med26 acts as a switch between preinitiation and elongation by interacting with components of TFIID during PIC formation, then exchanging those interactions for components of the super elongation complex (SEC), which includes pTEFb/cdk9 and the elongation factors ELL and EAF (319). Our data suggest that ICP4 may form transcription complexes with components of the Elongin SIII complex, not the SEC. Additionally, components of the 3'mRNA processing factor CPSF were isolated with wtICP4. Both of these complexes associate with the elongating form of RNA PolII (Phospho-Ser2 and Ser5). While RNA PolII is not phosphorylated on Ser2 throughout HSV infection, mechanisms to recruit the transcription elongation and mRNA processing machinery must be in place. It is possible that ICP4 may functionally replace Med26 and recruit elongation complexes (albeit different ones) and 3'mRNA processing complexes that normally require Ser2 phosphorylation.

Recruitment of 3' processing factors, in addition to components of Mediator and TFIID suggests a novel role in viral transcription. Cellular transcription has been shown to occur via a mechanism referred to as gene looping (292). Upon the onset of transcription initiation, many components of the transcription machinery remain at the promoter element, acting as a reinitiation scaffold. The reinitiation scaffold has been linked to the cleavage and polyadenylation specificity factor (CPSF) through interactions with TFIIB (301). These interactions form a "gene loop" that functions to enhance reinitiation of transcription and

establish “transcriptional memory”. Gene loops additionally enhance transcript export as they have been localized to regions near the nuclear pore (320). The findings herein support a model for viral transcription in which gene looping occurs. Work from the Knipe lab suggests that active transcription complexes form at the nuclear lamina (300). This implicates that these transcription complexes are at the periphery of the nucleus and may be in close proximity to nuclear pore complexes. The additional isolation of components of the export machinery with affinity purified ICP4 further supports this notion. Additionally, integrated HIV provirus forms gene loops in a TAT, TAR and CPSF dependent manner, providing an example of an RNA PolII dependent virus that utilize this mechanism to enhance viral transcription rates (256). Studies determining i.) gene loops are formed on the viral genome during HSV infection, ii.) the functional relevance of gene loops to HSV transcription, and iii.) the activities of ICP4 in their formation, would enhance the current understanding of transcription from DNA viruses. It is possible that this is a common mechanism utilized by many DNA viruses to increase the efficiency of viral transcription.

ICP4 Mediated Exit from Latency:

ICP4 has been shown to be important for viral replication in a mouse model of ocular infection. These studies indicated that regions contained within amino acids 142-210 were particularly important for viral replication in trigeminal ganglia, and as a consequence for efficient viral reactivation (11, 362). Interestingly, one mutant studied, ΔSER, showed no growth defect in cell culture yet failed to replicate *in vivo* (11). Some regions within the N-terminus have significant sequence homology between HSV-1, HSV-2, and Herpes B Virus, yet in cell culture, had no effect on viral replication when deleted either in the presence or absence of the C-terminal domain. This suggests that these regions either i. have redundant functions

encoded within the N-terminus, or ii. like ΔSER, have a significant role *in vivo*. Investigating the combined functions of these regions, specifically those deleted in ΔSER, m20, m90 and d143, would address whether these regions contain redundant functions in a cell culture model. The function of these regions in viral replication within the trigeminal ganglia, and the establishment and reactivation from latency are currently being investigated within the lab.

The differences observed in viral replication between the trigeminal ganglia and a cell culture model as a consequence of specific regions of ICP4 suggests that ICP4 has functions that are either redundant or dispensable within a cell culture model. The transcription environments in neurons and epithelial cells are likely different, implying that ICP4 may function differently based on the cellular transcription environment. The differences in transcriptional environments may arise partially from differing chromatin states of the genome. Studies presented herein demonstrated that ICP4 co-purified with ATP dependent chromatin remodeling complexes such as SWI/SNF, Ino80, and Nurd. It is possible that during the acute infection of the trigeminal ganglia, ICP4 mediated interactions with chromatin remodeling complexes may promote viral replication over the establishment of latency. Additionally, the potential for ICP4 to alter the chromatin conformation of the latent viral genome exists. In this respect, ICP0 would promote a general derepression of the viral genome (53, 100) and interactions between ICP4 and chromatin remodelers would promote efficient activation of viral transcription on a chromatinized template. In this manner, regions of ICP4 may be important for reactivation, which are otherwise dispensable for lytic infection. Studies to this affect are also ongoing within the lab.

The data presented in this thesis examined the structural and functional requirements of ICP4 mediated transcriptional activities. Together, these data advance the understanding of how ICP4 forms interactions with the transcription machinery and suggests new roles for ICP4 in

mediating viral transcription. Understanding the mechanisms underlying viral transcription can help advance fields aimed at developing novel drugs and vaccines against HSV as well as those aimed at utilizing HSV as a therapeutic agent.

6.0 BIBLIOGRAPHY

1. **Ahn, S. H., M. Kim, and S. Buratowski.** 2004. Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. Molecular cell **13:**67-76.
2. **Albright, S. R., and R. Tjian.** 2000. TAFs revisited: more data reveal new twists and confirm old ideas. Gene **242:**1-13.
3. **Alwine, J. C., W. L. Steinhart, and C. W. Hill.** 1974. Transcription of herpes simplex type 1 DNA in nuclei isolated from infected HEp-2 and KB cells. Virology **60:**302-307.
4. **Amelio, A. L., N. V. Giordani, N. J. Kubat, E. O'Neil J, and D. C. Bloom.** 2006. Deacetylation of the herpes simplex virus type 1 latency-associated transcript (LAT) enhancer and a decrease in LAT abundance precede an increase in ICP0 transcriptional permissiveness at early times postexplant. Journal of virology **80:**2063-2068.
5. **Amelio, A. L., P. K. McAnany, and D. C. Bloom.** 2006. A chromatin insulator-like element in the herpes simplex virus type 1 latency-associated transcript region binds CCCTC-binding factor and displays enhancer-blocking and silencing activities. Journal of virology **80:**2358-2368.
6. **Armache, K. J., H. Kettenberger, and P. Cramer.** 2003. Architecture of initiation-competent 12-subunit RNA polymerase II. Proceedings of the National Academy of Sciences of the United States of America **100:**6964-6968.
7. **Baek, H. J., Y. K. Kang, and R. G. Roeder.** 2006. Human Mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. The Journal of biological chemistry **281:**15172-15181.
8. **Baines, J. D., and S. G. Roberts.** 2011. Nuclear Egress and Envelopment of HSV, p. 195-206. *In* S. K. Weller (ed.), Alphaherpesviruses Molecular Virology. Caister Academic Press, Norfolk, UK.
9. **Balliet, J. W., and P. A. Schaffer.** 2006. Point mutations in herpes simplex virus type 1 oriL, but not in oriS, reduce pathogenesis during acute infection of mice and impair reactivation from latency. Journal of virology **80:**440-450.
10. **Barski, A., S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev, and K. Zhao.** 2007. High-resolution profiling of histone methylations in the human genome. Cell **129:**823-837.

11. **Bates, P. A., and N. A. DeLuca.** 1998. The polyserine tract of herpes simplex virus ICP4 is required for normal viral gene expression and growth in murine trigeminal ganglia. Journal of virology **72:**7115-7124.
12. **Batterson, W., D. Furlong, and B. Roizman.** 1983. Molecular genetics of herpes simplex virus. VIII. further characterization of a temperature-sensitive mutant defective in release of viral DNA and in other stages of the viral reproductive cycle. Journal of virology **45:**397-407.
13. **Batterson, W., and B. Roizman.** 1983. Characterization of the herpes simplex virion-associated factor responsible for the induction of alpha genes. Journal of virology **46:**371-377.
14. **Beard, P. M., N. S. Taus, and J. D. Baines.** 2002. DNA cleavage and packaging proteins encoded by genes U(L)28, U(L)15, and U(L)33 of herpes simplex virus type 1 form a complex in infected cells. Journal of virology **76:**4785-4791.
15. **Belotserkovskaya, R., A. Saunders, J. T. Lis, and D. Reinberg.** 2004. Transcription through chromatin: understanding a complex FACT. Biochimica et biophysica acta **1677:**87-99.
16. **Benyesh-Melnick, M., P. A. Schaffer, R. J. Courtney, J. Esparza, and S. Kimura.** 1975. Viral gene functions expressed and detected by temperature-sensitive mutants of herpes simplex virus. Cold Spring Harbor symposia on quantitative biology **39 Pt 2:**731-746.
17. **Bertke, A. S., A. Patel, Y. Imai, K. Apakupakul, T. P. Margolis, and P. R. Krause.** 2009. Latency-associated transcript (LAT) exon 1 controls herpes simplex virus species-specific phenotypes: reactivation in the guinea pig genital model and neuron subtype-specific latent expression of LAT. Journal of virology **83:**10007-10015.
18. **Bhaumik, S. R.** 2011. Distinct regulatory mechanisms of eukaryotic transcriptional activation by SAGA and TFIID. Biochimica et biophysica acta **1809:**97-108.
19. **Blaho, J. A., N. Michael, V. Kang, N. Aboul-Ela, M. E. Smulson, M. K. Jacobson, and B. Roizman.** 1992. Differences in the poly(ADP-ribosyl)ation patterns of ICP4, the herpes simplex virus major regulatory protein, in infected cells and in isolated nuclei. Journal of virology **66:**6398-6407.
20. **Blaho, J. A., and B. Roizman.** 1991. ICP4, the major regulatory protein of herpes simplex virus, shares features common to GTP-binding proteins and is adenylated and guanylated. Journal of virology **65:**3759-3769.
21. **Bloom, D. C., N. V. Giordani, and D. L. Kwiatkowski.** 2010. Epigenetic regulation of latent HSV-1 gene expression. Biochimica et biophysica acta **1799:**246-256.
22. **Bloom, D. C., J. M. Hill, G. Devi-Rao, E. K. Wagner, L. T. Feldman, and J. G. Stevens.** 1996. A 348-base-pair region in the latency-associated transcript facilitates herpes simplex virus type 1 reactivation. Journal of virology **70:**2449-2459.
23. **Blumel, J., and B. Matz.** 1995. Thermosensitive UL9 gene function is required for early stages of herpes simplex virus type 1 DNA synthesis. The Journal of general virology **76 (Pt 12):**3119-3124.
24. **Boehmer, P. E.** 1998. The herpes simplex virus type-1 single-strand DNA-binding protein, ICP8, increases the processivity of the UL9 protein DNA helicase. The Journal of biological chemistry **273:**2676-2683.

25. **Booy, F. P., W. W. Newcomb, B. L. Trus, J. C. Brown, T. S. Baker, and A. C. Steven.** 1991. Liquid-crystalline, phage-like packing of encapsidated DNA in herpes simplex virus. Cell **64:**1007-1015.
26. **Bradsher, J. N., K. W. Jackson, R. C. Conaway, and J. W. Conaway.** 1993. RNA polymerase II transcription factor SIII. I. Identification, purification, and properties. The Journal of biological chemistry **268:**25587-25593.
27. **Bradsher, J. N., S. Tan, H. J. McLaury, J. W. Conaway, and R. C. Conaway.** 1993. RNA polymerase II transcription factor SIII. II. Functional properties and role in RNA chain elongation. The Journal of biological chemistry **268:**25594-25603.
28. **Brugha, R., K. Keersmaekers, A. Renton, and A. Meheus.** 1997. Genital herpes infection: a review. International journal of epidemiology **26:**698-709.
29. **Bryant, H. E., S. E. Wadd, A. I. Lamond, S. J. Silverstein, and J. B. Clements.** 2001. Herpes simplex virus IE63 (ICP27) protein interacts with spliceosome-associated protein 145 and inhibits splicing prior to the first catalytic step. Journal of virology **75:**4376-4385.
30. **Bryant, K. F., R. C. Colgrove, and D. M. Knipe.** 2011. Cellular SNF2H chromatin-remodeling factor promotes herpes simplex virus 1 immediate-early gene expression and replication. mBio **2:**e00330-00310.
31. **Buratowski, S.** 2009. Progression through the RNA polymerase II CTD cycle. Molecular cell **36:**541-546.
32. **Buratowski, S., S. Hahn, L. Guarente, and P. A. Sharp.** 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. Cell **56:**549-561.
33. **Burke, T. W., and J. T. Kadonaga.** 1997. The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila. Genes & development **11:**3020-3031.
34. **Burke, T. W., and J. T. Kadonaga.** 1996. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes & development **10:**711-724.
35. **Bussmann, C., W. M. Peng, T. Bieber, and N. Novak.** 2008. Molecular pathogenesis and clinical implications of eczema herpeticum. Expert reviews in molecular medicine **10:**e21.
36. **Calvo, O., and J. L. Manley.** 2005. The transcriptional coactivator PC4/Sub1 has multiple functions in RNA polymerase II transcription. The EMBO journal **24:**1009-1020.
37. **Campadelli-Fiume, G., C. De Giovanni, V. Gatta, P. Nanni, P. L. Lollini, and L. Menotti.** 2011. Rethinking herpes simplex virus: the way to oncolytic agents. Reviews in medical virology **21:**213-226.
38. **Campbell, M. E., J. W. Palfreyman, and C. M. Preston.** 1984. Identification of herpes simplex virus DNA sequences which encode a trans-acting polypeptide responsible for stimulation of immediate early transcription. Journal of molecular biology **180:**1-19.
39. **Campos, E. I., and D. Reinberg.** 2009. Histones: annotating chromatin. Annual review of genetics **43:**559-599.
40. **Carlberg, C., and S. Seuter.** 2010. Dynamics of nuclear receptor target gene regulation. Chromosoma **119:**479-484.
41. **Carrozza, M. J., and N. A. DeLuca.** 1996. Interaction of the viral activator protein ICP4 with TFIID through TAF250. Molecular and cellular biology **16:**3085-3093.

42. **Carter, C. J.** 2008. Interactions between the products of the Herpes simplex genome and Alzheimer's disease susceptibility genes: relevance to pathological-signalling cascades. Neurochemistry international **52:**920-934.
43. **Chalkley, G. E., and C. P. Verrijzer.** 1999. DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. The EMBO journal **18:**4835-4845.
44. **Chen, H. T., and S. Hahn.** 2003. Binding of TFIIB to RNA polymerase II: Mapping the binding site for the TFIIB zinc ribbon domain within the preinitiation complex. Molecular cell **12:**437-447.
45. **Chen, I. H., L. Li, L. Silva, and R. M. Sandri-Goldin.** 2005. ICP27 recruits Aly/REF but not TAP/NXF1 to herpes simplex virus type 1 transcription sites although TAP/NXF1 is required for ICP27 export. Journal of virology **79:**3949-3961.
46. **Chen, I. H., K. S. Sciabica, and R. M. Sandri-Goldin.** 2002. ICP27 interacts with the RNA export factor Aly/REF to direct herpes simplex virus type 1 intronless mRNAs to the TAP export pathway. Journal of virology **76:**12877-12889.
47. **Clapier, C. R., and B. R. Cairns.** 2009. The biology of chromatin remodeling complexes. Annual review of biochemistry **78:**273-304.
48. **Cliffe, A. R., D. A. Garber, and D. M. Knipe.** 2009. Transcription of the herpes simplex virus latency-associated transcript promotes the formation of facultative heterochromatin on lytic promoters. Journal of virology **83:**8182-8190.
49. **Cliffe, A. R., and D. M. Knipe.** 2008. Herpes simplex virus ICP0 promotes both histone removal and acetylation on viral DNA during lytic infection. Journal of virology **82:**12030-12038.
50. **Cockrell, S. K.** 2010. The Role of The Herpes Simplex Virus Type 1 UL25 Protein in DNA Packaging and Virion Assembly. The University of Pittsburgh, Pittsburgh, PA.
51. **Coen, D. M., and P. A. Schaffer.** 1980. Two distinct loci confer resistance to acycloguanosine in herpes simplex virus type 1. Proceedings of the National Academy of Sciences of the United States of America **77:**2265-2269.
52. **Coen, D. M., S. P. Weinheimer, and S. L. McKnight.** 1986. A genetic approach to promoter recognition during trans induction of viral gene expression. Science **234:**53-59.
53. **Coleman, H. M., V. Connor, Z. S. Cheng, F. Grey, C. M. Preston, and S. Efstathiou.** 2008. Histone modifications associated with herpes simplex virus type 1 genomes during quiescence and following ICP0-mediated de-repression. The Journal of general virology **89:**68-77.
54. **Compel, P., and N. A. DeLuca.** 2003. Temperature-dependent conformational changes in herpes simplex virus ICP4 that affect transcription activation. Journal of virology **77:**3257-3268.
55. **Conaway, J. W., and R. C. Conaway.** 1989. A multisubunit transcription factor essential for accurate initiation by RNA polymerase II. The Journal of biological chemistry **264:**2357-2362.
56. **Conaway, R. C., and J. W. Conaway.** 2009. The INO80 chromatin remodeling complex in transcription, replication and repair. Trends in biochemical sciences **34:**71-77.
57. **Conaway, R. C., and J. W. Conaway.** 1989. An RNA polymerase II transcription factor has an associated DNA-dependent ATPase (dATPase) activity strongly stimulated by the TATA region of promoters. Proceedings of the National Academy of Sciences of the United States of America **86:**7356-7360.

58. **Conway, J. F., S. K. Cockrell, A. M. Copeland, W. W. Newcomb, J. C. Brown, and F. L. Homa.** 2010. Labeling and localization of the herpes simplex virus capsid protein UL25 and its interaction with the two triplexes closest to the penton. Journal of molecular biology **397:**575-586.
59. **Conway, J. F., and F. L. Homa.** 2011. Nucleocapsid Structure, Assemby and DNA Packaging of Herpes Simplex Virus, p. 175-194. *In* S. K. Weller (ed.), Alphaherpesviruses Molecular Virology. Caister Academic Press, Norfolk, UK.
60. **Courtney, R. J., and M. Benyesh-Melnick.** 1974. Isolation and characterization of a large molecular-weight polypeptide of herpes simplex virus type 1. Virology **62:**539-551.
61. **Cui, C., A. Griffiths, G. Li, L. M. Silva, M. F. Kramer, T. Gaasterland, X. J. Wang, and D. M. Coen.** 2006. Prediction and identification of herpes simplex virus 1-encoded microRNAs. Journal of virology **80:**5499-5508.
62. **Curanovic, D., and L. Enquist.** 2009. Directional transneuronal spread of alpha-herpesvirus infection. Future virology **4:**591.
63. **Dai-Ju, J. Q., L. Li, L. A. Johnson, and R. M. Sandri-Goldin.** 2006. ICP27 interacts with the C-terminal domain of RNA polymerase II and facilitates its recruitment to herpes simplex virus 1 transcription sites, where it undergoes proteasomal degradation during infection. Journal of virology **80:**3567-3581.
64. **Delboy, M. G., J. L. Patterson, A. M. Hollander, and A. V. Nicola.** 2006. Nectin-2-mediated entry of a syncytial strain of herpes simplex virus via pH-independent fusion with the plasma membrane of Chinese hamster ovary cells. Virology journal **3:**105.
65. **DeLuca, N.** 2011. Functions and Mechanism of Action of the Herpes Simplex Virus Regulatory Protein, ICP4, p. 17-38. *In* S. K. Weller (ed.), Alphaherpesviruses Molecular Virology. Caister Academic Press, Norfolk, UK.
66. **DeLuca, N.** 2011. Functionsand Mechanism of Action of the Herpes Simplex Virus Regulatory Protein, ICP4, p. 17-38. *In* S. K. Weller (ed.), Alphaherpesviruses Molecular Virology. Caister Academic Press, Norfolk, UK.
67. **DeLuca, N. A., A. M. McCarthy, and P. A. Schaffer.** 1985. Isolation and characterization of deletion mutants of herpes simplex virus type 1 in the gene encoding immediate-early regulatory protein ICP4. Journal of virology **56:**558-570.
68. **DeLuca, N. A., and P. A. Schaffer.** 1985. Activation of immediate-early, early, and late promoters by temperature-sensitive and wild-type forms of herpes simplex virus type 1 protein ICP4. Molecular and cellular biology **5:**1997-2008.
69. **DeLuca, N. A., and P. A. Schaffer.** 1987. Activities of herpes simplex virus type 1 (HSV-1) ICP4 genes specifying nonsense peptides. Nucleic acids research **15:**4491-4511.
70. **DeLuca, N. A., and P. A. Schaffer.** 1988. Physical and functional domains of the herpes simplex virus transcriptional regulatory protein ICP4. Journal of virology **62:**732-743.
71. **Deshmane, S. L., and N. W. Fraser.** 1989. During latency, herpes simplex virus type 1 DNA is associated with nucleosomes in a chromatin structure. Journal of virology **63:**943-947.
72. **Devi-Rao, G. B., D. C. Bloom, J. G. Stevens, and E. K. Wagner.** 1994. Herpes simplex virus type 1 DNA replication and gene expression during explant-induced reactivation of latently infected murine sensory ganglia. Journal of virology **68:**1271-1282.
73. **Dhar, S., and J. P. Weir.** 2000. Herpes simplex virus 1 late gene expression is preferentially inhibited during infection of the TAF250 mutant ts13 cell line. Virology **270:**190-200.

74. **DiDonato, J. A., and M. T. Muller.** 1989. DNA binding and gene regulation by the herpes simplex virus type 1 protein ICP4 and involvement of the TATA element. Journal of virology **63:**3737-3747.
75. **DiDonato, J. A., J. R. Spitzner, and M. T. Muller.** 1991. A predictive model for DNA recognition by the herpes simplex virus protein ICP4. Journal of molecular biology **219:**451-470.
76. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic acids research **11:**1475-1489.
77. **Dixon, R. A., and P. A. Schaffer.** 1980. Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175. Journal of virology **36:**189-203.
78. **Dohner, K., A. Wolfstein, U. Prank, C. Echeverri, D. Dujardin, R. Vallee, and B. Sodeik.** 2002. Function of dynein and dynactin in herpes simplex virus capsid transport. Molecular biology of the cell **13:**2795-2809.
79. **Dominski, Z.** 2010. The hunt for the 3' endonuclease. Wiley interdisciplinary reviews. RNA **1:**325-340.
80. **Douglas, M. W., R. J. Diefenbach, F. L. Homa, M. Miranda-Saksena, F. J. Rixon, V. Vittone, K. Byth, and A. L. Cunningham.** 2004. Herpes simplex virus type 1 capsid protein VP26 interacts with dynein light chains RP3 and Tctex1 and plays a role in retrograde cellular transport. The Journal of biological chemistry **279:**28522-28530.
81. **Dracheva, S., E. V. Koonin, and J. J. Crute.** 1995. Identification of the primase active site of the herpes simplex virus type 1 helicase-primase. The Journal of biological chemistry **270:**14148-14153.
82. **Durand, L. O., and B. Roizman.** 2008. Role of cdk9 in the optimization of expression of the genes regulated by ICP22 of herpes simplex virus 1. Journal of virology **82:**10591-10599.
83. **Dyson, H. J., and P. E. Wright.** 2005. Intrinsically unstructured proteins and their functions. Nature reviews. Molecular cell biology **6:**197-208.
84. **Egloff, S., M. Dienstbier, and S. Murphy.** 2012. Updating the RNA polymerase CTD code: adding gene-specific layers. Trends in genetics : TIG.
85. **Egly, J. M., and F. Coin.** 2011. A history of TFIIH: two decades of molecular biology on a pivotal transcription/repair factor. DNA repair **10:**714-721.
86. **Eisenberg, R. J., E. E. Heldwein, G. H. Cohen, and C. Krummenacher.** 2011. Recent Progress in Understanding Herpes Simplex Virus Entry: Relationship of Structure and Function, p. 131-152. *In* S. K. Weller (ed.), Alphaherpesviruses Molecular Virology. Caister Academic Press, Norfolk, UK.
87. **Eisenberg, S. P., D. M. Coen, and S. L. McKnight.** 1985. Promoter domains required for expression of plasmid-borne copies of the herpes simplex virus thymidine kinase gene in virus-infected mouse fibroblasts and microinjected frog oocytes. Molecular and cellular biology **5:**1940-1947.
88. **Elion, G. B.** 1982. Mechanism of action and selectivity of acyclovir. The American journal of medicine **73:**7-13.
89. **Elion, G. B., P. A. Furman, J. A. Fyfe, P. de Miranda, L. Beauchamp, and H. J. Schaeffer.** 1977. Selectivity of action of an antiherpetic agent, 9-(2-

hydroxyethoxymethyl) guanine. Proceedings of the National Academy of Sciences of the United States of America **74:**5716-5720.

90. **Everett, R. D.** 1984. A detailed analysis of an HSV-1 early promoter: sequences involved in trans-activation by viral immediate-early gene products are not early-gene specific. Nucleic acids research **12:**3037-3056.
91. **Everett, R. D., C. Parada, P. Gripon, H. Sirma, and A. Orr.** 2008. Replication of ICP0-null mutant herpes simplex virus type 1 is restricted by both PML and Sp100. Journal of virology **82:**2661-2672.
92. **Everett, R. D., G. Sourvinos, and A. Orr.** 2003. Recruitment of herpes simplex virus type 1 transcriptional regulatory protein ICP4 into foci juxtaposed to ND10 in live, infected cells. Journal of virology **77:**3680-3689.
93. **Faber, S. W., and K. W. Wilcox.** 1988. Association of herpes simplex virus regulatory protein ICP4 with sequences spanning the ICP4 gene transcription initiation site. Nucleic acids research **16:**555-570.
94. **Faber, S. W., and K. W. Wilcox.** 1986. Association of the herpes simplex virus regulatory protein ICP4 with specific nucleotide sequences in DNA. Nucleic acids research **14:**6067-6083.
95. **Faber, S. W., and K. W. Wilcox.** 1986. Characterization of a herpes simplex virus regulatory protein: aggregation and phosphorylation of a temperature-sensitive variant of ICP 4. Archives of virology **91:**297-312.
96. **Fan, X., D. M. Chou, and K. Struhl.** 2006. Activator-specific recruitment of Mediator in vivo. Nature structural & molecular biology **13:**117-120.
97. **Farrell, M. J., T. P. Margolis, W. A. Gomes, and L. T. Feldman.** 1994. Effect of the transcription start region of the herpes simplex virus type 1 latency-associated transcript promoter on expression of productively infected neurons in vivo. Journal of virology **68:**5337-5343.
98. **Felser, J. M., P. R. Kinchington, G. Inchauspe, S. E. Straus, and J. M. Ostrove.** 1988. Cell lines containing varicella-zoster virus open reading frame 62 and expressing the "IE" 175 protein complement ICP4 mutants of herpes simplex virus type 1. Journal of virology **62:**2076-2082.
99. **Ferenczy, M. W., and N. A. DeLuca.** 2009. Epigenetic modulation of gene expression from quiescent herpes simplex virus genomes. Journal of virology **83:**8514-8524.
100. **Ferenczy, M. W., and N. A. DeLuca.** 2011. Reversal of heterochromatic silencing of quiescent herpes simplex virus type 1 by ICP0. Journal of virology **85:**3424-3435.
101. **Ferenczy, M. W., D. J. Ranayhossaini, and N. A. Deluca.** 2011. Activities of ICP0 involved in the reversal of silencing of quiescent herpes simplex virus 1. Journal of virology **85:**4993-5002.
102. **Field, H. J., G. Darby, and P. Wildy.** 1980. Isolation and characterization of acyclovir-resistant mutants of herpes simplex virus. The Journal of general virology **49:**115-124.
103. **Fink, D. J., J. Wechuck, M. Mata, J. C. Glorioso, J. Goss, D. Krisky, and D. Wolfe.** 2011. Gene therapy for pain: results of a phase I clinical trial. Annals of neurology **70:**207-212.
104. **Fire, A., M. Samuels, and P. A. Sharp.** 1984. Interactions between RNA polymerase II, factors, and template leading to accurate transcription. The Journal of biological chemistry **259:**2509-2516.
105. **Flint, S. J.** 2004. Principles of Virology, 2nd ed. ASM press, Washington DC.

106. **Flores, O., H. Lu, M. Killeen, J. Greenblatt, Z. F. Burton, and D. Reinberg.** 1991. The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. Proceedings of the National Academy of Sciences of the United States of America **88:**9999-10003.
107. **Flores, O., H. Lu, and D. Reinberg.** 1992. Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIH. The Journal of biological chemistry **267:**2786-2793.
108. **Fraser, K. A., and S. A. Rice.** 2005. Herpes simplex virus type 1 infection leads to loss of serine-2 phosphorylation on the carboxyl-terminal domain of RNA polymerase II. Journal of virology **79:**11323-11334.
109. **Freeman, E. E., H. A. Weiss, J. R. Glynn, P. L. Cross, J. A. Whitworth, and R. J. Hayes.** 2006. Herpes simplex virus 2 infection increases HIV acquisition in men and women: systematic review and meta-analysis of longitudinal studies. AIDS **20:**73-83.
110. **Ge, H., and R. G. Roeder.** 1994. Purification, cloning, and characterization of a human coactivator, PC4, that mediates transcriptional activation of class II genes. Cell **78:**513-523.
111. **Gerard, M., L. Fischer, V. Moncollin, J. M. Chipoulet, P. Chambon, and J. M. Egly.** 1991. Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2. The Journal of biological chemistry **266:**20940-20945.
112. **Ghazy, M. A., S. A. Brodie, M. L. Ammerman, L. M. Ziegler, and A. S. Ponticelli.** 2004. Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. Molecular and cellular biology **24:**10975-10985.
113. **Gibson, W., and B. Roizman.** 1972. Proteins specified by herpes simplex virus. 8. Characterization and composition of multiple capsid forms of subtypes 1 and 2. Journal of virology **10:**1044-1052.
114. **Godowski, P. J., and D. M. Knipe.** 1986. Transcriptional control of herpesvirus gene expression: gene functions required for positive and negative regulation. Proceedings of the National Academy of Sciences of the United States of America **83:**256-260.
115. **Gottlieb, J., A. I. Marcy, D. M. Coen, and M. D. Challberg.** 1990. The herpes simplex virus type 1 UL42 gene product: a subunit of DNA polymerase that functions to increase processivity. Journal of virology **64:**5976-5987.
116. **Grewal, S. I., and S. Jia.** 2007. Heterochromatin revisited. Nature reviews. Genetics **8:**35-46.
117. **Grondin, B., and N. DeLuca.** 2000. Herpes simplex virus type 1 ICP4 promotes transcription preinitiation complex formation by enhancing the binding of TFIID to DNA. Journal of virology **74:**11504-11510.
118. **Gu, B., and N. DeLuca.** 1994. Requirements for activation of the herpes simplex virus glycoprotein C promoter in vitro by the viral regulatory protein ICP4. Journal of virology **68:**7953-7965.
119. **Gu, B., R. Kuddus, and N. A. DeLuca.** 1995. Repression of activator-mediated transcription by herpes simplex virus ICP4 via a mechanism involving interactions with the basal transcription factors TATA-binding protein and TFIIB. Molecular and cellular biology **15:**3618-3626.
120. **Gu, B., R. Rivera-Gonzalez, C. A. Smith, and N. A. DeLuca.** 1993. Herpes simplex virus infected cell polypeptide 4 preferentially represses Sp1-activated over basal

transcription from its own promoter. Proceedings of the National Academy of Sciences of the United States of America **90:**9528-9532.

121. **Gu, H., and B. Roizman.** 2007. Herpes simplex virus-infected cell protein 0 blocks the silencing of viral DNA by dissociating histone deacetylases from the CoREST-REST complex. Proceedings of the National Academy of Sciences of the United States of America **104:**17134-17139.
122. **Ha, I., S. Roberts, E. Maldonado, X. Sun, L. U. Kim, M. Green, and D. Reinberg.** 1993. Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. Genes & development **7:**1021-1032.
123. **Hahn, S., and E. T. Young.** 2011. Transcriptional regulation in Saccharomyces cerevisiae: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. Genetics **189:**705-736.
124. **Halford, W. P., C. D. Kemp, J. A. Isler, D. J. Davido, and P. A. Schaffer.** 2001. ICP0, ICP4, or VP16 expressed from adenovirus vectors induces reactivation of latent herpes simplex virus type 1 in primary cultures of latently infected trigeminal ganglion cells. Journal of virology **75:**6143-6153.
125. **Halpern, M. E., and J. R. Smiley.** 1984. Effects of deletions on expression of the herpes simplex virus thymidine kinase gene from the intact viral genome: the amino terminus of the enzyme is dispensable for catalytic activity. Journal of virology **50:**733-738.
126. **Hargreaves, D. C., and G. R. Crabtree.** 2011. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. Cell research **21:**396-420.
127. **Hayward, G. S., R. J. Jacob, S. C. Wadsworth, and B. Roizman.** 1975. Anatomy of herpes simplex virus DNA: evidence for four populations of molecules that differ in the relative orientations of their long and short components. Proceedings of the National Academy of Sciences of the United States of America **72:**4243-4247.
128. **Heldwein, E. E., and C. Krummenacher.** 2008. Entry of herpesviruses into mammalian cells. Cellular and molecular life sciences : CMLS **65:**1653-1668.
129. **Herrera, F. J., and S. J. Triezenberg.** 2004. VP16-dependent association of chromatin-modifying coactivators and underrepresentation of histones at immediate-early gene promoters during herpes simplex virus infection. Journal of virology **78:**9689-9696.
130. **Hill, J. M., F. Sedarati, R. T. Javier, E. K. Wagner, and J. G. Stevens.** 1990. Herpes simplex virus latent phase transcription facilitates in vivo reactivation. Virology **174:**117-125.
131. **Hodge, P. D., and N. D. Stow.** 2001. Effects of mutations within the herpes simplex virus type 1 DNA encapsidation signal on packaging efficiency. Journal of virology **75:**8977-8986.
132. **Holland, D. J., M. Miranda-Saksena, R. A. Boadle, P. Armati, and A. L. Cunningham.** 1999. Anterograde transport of herpes simplex virus proteins in axons of peripheral human fetal neurons: an immunoelectron microscopy study. Journal of virology **73:**8503-8511.
133. **Holstege, F. C., P. C. van der Vliet, and H. T. Timmers.** 1996. Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. The EMBO journal **15:**1666-1677.
134. **Homa, F. L., and J. C. Brown.** 1997. Capsid assembly and DNA packaging in herpes simplex virus. Reviews in medical virology **7:**107-122.

135. **Honess, R. W., and B. Roizman.** 1975. Regulation of herpesvirus macromolecular synthesis: sequential transition of polypeptide synthesis requires functional viral polypeptides. Proceedings of the National Academy of Sciences of the United States of America **72:**1276-1280.
136. **Honess, R. W., and B. Roizman.** 1974. Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. Journal of virology **14:**8-19.
137. **Hoshino, Y., S. K. Dalai, K. Wang, L. Pesnicak, T. Y. Lau, D. M. Knipe, J. I. Cohen, and S. E. Straus.** 2005. Comparative efficacy and immunogenicity of replication-defective, recombinant glycoprotein, and DNA vaccines for herpes simplex virus 2 infections in mice and guinea pigs. Journal of virology **79:**410-418.
138. **Huang, C. J., and E. K. Wagner.** 1994. The herpes simplex virus type 1 major capsid protein (VP5-UL19) promoter contains two cis-acting elements influencing late expression. Journal of virology **68:**5738-5747.
139. **Huang, J., J. R. Kent, B. Placek, K. A. Whelan, C. M. Hollow, P. Y. Zeng, N. W. Fraser, and S. L. Berger.** 2006. Trimethylation of histone H3 lysine 4 by Set1 in the lytic infection of human herpes simplex virus 1. Journal of virology **80:**5740-5746.
140. **Imbalzano, A. N., D. M. Coen, and N. A. DeLuca.** 1991. Herpes simplex virus transactivator ICP4 operationally substitutes for the cellular transcription factor Sp1 for efficient expression of the viral thymidine kinase gene. Journal of virology **65:**565-574.
141. **Imbalzano, A. N., and N. A. DeLuca.** 1992. Substitution of a TATA box from a herpes simplex virus late gene in the viral thymidine kinase promoter alters ICP4 inducibility but not temporal expression. Journal of virology **66:**5453-5463.
142. **Ishov, A. M., and G. G. Maul.** 1996. The periphery of nuclear domain 10 (ND10) as site of DNA virus deposition. The Journal of cell biology **134:**815-826.
143. **Ito, M., C. X. Yuan, S. Malik, W. Gu, J. D. Fondell, S. Yamamura, Z. Y. Fu, X. Zhang, J. Qin, and R. G. Roeder.** 1999. Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators. Molecular cell **3:**361-370.
144. **Iwahori, S., N. Shirata, Y. Kawaguchi, S. K. Weller, Y. Sato, A. Kudoh, S. Nakayama, H. Isomura, and T. Tsurumi.** 2007. Enhanced phosphorylation of transcription factor sp1 in response to herpes simplex virus type 1 infection is dependent on the ataxia telangiectasia-mutated protein. Journal of virology **81:**9653-9664.
145. **Jackson, S. A., and N. A. DeLuca.** 2003. Relationship of herpes simplex virus genome configuration to productive and persistent infections. Proceedings of the National Academy of Sciences of the United States of America **100:**7871-7876.
146. **Jacob, R. J., L. S. Morse, and B. Roizman.** 1979. Anatomy of herpes simplex virus DNA. XII. Accumulation of head-to-tail concatemers in nuclei of infected cells and their role in the generation of the four isomeric arrangements of viral DNA. Journal of virology **29:**448-457.
147. **Jenuwein, T., and C. D. Allis.** 2001. Translating the histone code. Science **293:**1074-1080.
148. **Jessie K., H., O.H., Rahim, Z.H.A.** 2008. Protein purification method for salivary proteins and rehydration buffer for two dimensional electrophoresis. Biotechnology **7:**686-693.

149. **Jiang, C., and B. F. Pugh.** 2009. A compiled and systematic reference map of nucleosome positions across the Saccharomyces cerevisiae genome. Genome biology **10:**R109.
150. **Jiang, C., and B. F. Pugh.** 2009. Nucleosome positioning and gene regulation: advances through genomics. Nature reviews. Genetics **10:**161-172.
151. **Johnson, L. A., L. Li, and R. M. Sandri-Goldin.** 2009. The cellular RNA export receptor TAP/NXF1 is required for ICP27-mediated export of herpes simplex virus 1 RNA, but the TREX complex adaptor protein Aly/REF appears to be dispensable. Journal of virology **83:**6335-6346.
152. **Johnson, L. A., and R. M. Sandri-Goldin.** 2009. Efficient nuclear export of herpes simplex virus 1 transcripts requires both RNA binding by ICP27 and ICP27 interaction with TAP/NXF1. Journal of virology **83:**1184-1192.
153. **Johnson, R.** 2004. Herpes gladiatorum and other skin diseases. Clinics in sports medicine **23:**473-484, x.
154. **Johnson, T. L., and J. Vilardell.** 2012. Regulated pre-mRNA splicing: The ghostwriter of the eukaryotic genome. Biochimica et biophysica acta **1819:**538-545.
155. **Jones, K. A., K. R. Yamamoto, and R. Tjian.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. Cell **42:**559-572.
156. **Juven-Gershon, T., J. Y. Hsu, J. W. Theisen, and J. T. Kadonaga.** 2008. The RNA polymerase II core promoter - the gateway to transcription. Current opinion in cell biology **20:**253-259.
157. **Kang, J. J., D. T. Auble, J. A. Ranish, and S. Hahn.** 1995. Analysis of the yeast transcription factor TFIIA: distinct functional regions and a polymerase II-specific role in basal and activated transcription. Molecular and cellular biology **15:**1234-1243.
158. **Katahira, J.** 2012. mRNA export and the TREX complex. Biochimica et biophysica acta **1819:**507-513.
159. **Kato, R., D. Wolfe, C. H. Coyle, J. B. Wechuck, P. Tyagi, T. Tsukamoto, J. B. Nelson, J. C. Glorioso, M. B. Chancellor, and N. Yoshimura.** 2009. Herpes simplex virus vector-mediated delivery of neurturin rescues erectile dysfunction of cavernous nerve injury. Gene therapy **16:**26-33.
160. **Kaye, S., and A. Choudhary.** 2006. Herpes simplex keratitis. Progress in retinal and eye research **25:**355-380.
161. **Kent, J. R., P. Y. Zeng, D. Atanasiu, J. Gardner, N. W. Fraser, and S. L. Berger.** 2004. During lytic infection herpes simplex virus type 1 is associated with histones bearing modifications that correlate with active transcription. Journal of virology **78:**10178-10186.
162. **Kim, D. B., and N. A. DeLuca.** 2002. Phosphorylation of transcription factor Sp1 during herpes simplex virus type 1 infection. Journal of virology **76:**6473-6479.
163. **Kim, D. B., S. Zabierowski, and N. A. DeLuca.** 2002. The initiator element in a herpes simplex virus type 1 late-gene promoter enhances activation by ICP4, resulting in abundant late-gene expression. Journal of virology **76:**1548-1558.
164. **Kim, J. Y., A. Mandarino, M. V. Chao, I. Mohr, and A. C. Wilson.** 2012. Transient reversal of episome silencing precedes VP16-dependent transcription during reactivation of latent HSV-1 in neurons. PLoS pathogens **8:**e1002540.

165. **Kim, M., H. Suh, E. J. Cho, and S. Buratowski.** 2009. Phosphorylation of the yeast Rpb1 C-terminal domain at serines 2, 5, and 7. The Journal of biological chemistry **284:**26421-26426.
166. **King, S. M.** 2000. The dynein microtubule motor. Biochimica et biophysica acta **1496:**60-75.
167. **Knipe, D. M., and A. Cliffe.** 2008. Chromatin control of herpes simplex virus lytic and latent infection. Nature reviews. Microbiology **6:**211-221.
168. **Knipe, D. M., W. T. Ruyechan, B. Roizman, and I. W. Halliburton.** 1978. Molecular genetics of herpes simplex virus: demonstration of regions of obligatory and nonobligatory identity within diploid regions of the genome by sequence replacement and insertion. Proceedings of the National Academy of Sciences of the United States of America **75:**3896-3900.
169. **Knipe, D. M., D. Senechek, S. A. Rice, and J. L. Smith.** 1987. Stages in the nuclear association of the herpes simplex virus transcriptional activator protein ICP4. Journal of virology **61:**276-284.
170. **Kobayashi, N., T. G. Boyer, and A. J. Berk.** 1995. A class of activation domains interacts directly with TFIIA and stimulates TFIIA-TFIID-promoter complex assembly. Molecular and cellular biology **15:**6465-6473.
171. **Koff, A., J. F. Schwedes, and P. Tegtmeyer.** 1991. Herpes simplex virus origin-binding protein (UL9) loops and distorts the viral replication origin. Journal of virology **65:**3284-3292.
172. **Komarnitsky, P., E. J. Cho, and S. Buratowski.** 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes & development **14:**2452-2460.
173. **Kornberg, R. D.** 1974. Chromatin structure: a repeating unit of histones and DNA. Science **184:**868-871.
174. **Kornberg, R. D.** 2005. Mediator and the mechanism of transcriptional activation. Trends in biochemical sciences **30:**235-239.
175. **Krishnamurthy, M., A. Dugan, A. Nwokoye, Y. H. Fung, J. K. Lancia, C. Y. Majmudar, and A. K. Mapp.** 2011. Caught in the act: covalent cross-linking captures activator-coactivator interactions in vivo. ACS chemical biology **6:**1321-1326.
176. **Kristie, T. M., and B. Roizman.** 1986. Alpha 4, the major regulatory protein of herpes simplex virus type 1, is stably and specifically associated with promoter-regulatory domains of alpha genes and of selected other viral genes. Proceedings of the National Academy of Sciences of the United States of America **83:**3218-3222.
177. **Kristie, T. M., and B. Roizman.** 1986. DNA-binding site of major regulatory protein alpha 4 specifically associated with promoter-regulatory domains of alpha genes of herpes simplex virus type 1. Proceedings of the National Academy of Sciences of the United States of America **83:**4700-4704.
178. **Kristie, T. M., and P. A. Sharp.** 1990. Interactions of the Oct-1 POU subdomains with specific DNA sequences and with the HSV alpha-trans-activator protein. Genes & development **4:**2383-2396.
179. **Krummenacher, C., A. V. Nicola, J. C. Whitbeck, H. Lou, W. Hou, J. D. Lambris, R. J. Geraghty, P. G. Spear, G. H. Cohen, and R. J. Eisenberg.** 1998. Herpes simplex virus glycoprotein D can bind to poliovirus receptor-related protein 1 or herpesvirus entry

mediator, two structurally unrelated mediators of virus entry. Journal of virology **72:**7064-7074.

180. **Kubat, N. J., A. L. Amelio, N. V. Giordani, and D. C. Bloom.** 2004. The herpes simplex virus type 1 latency-associated transcript (LAT) enhancer/rcr is hyperacetylated during latency independently of LAT transcription. Journal of virology **78:**12508-12518.
181. **Kubat, N. J., R. K. Tran, P. McAnany, and D. C. Bloom.** 2004. Specific histone tail modification and not DNA methylation is a determinant of herpes simplex virus type 1 latent gene expression. Journal of virology **78:**1139-1149.
182. **Kuddus, R., B. Gu, and N. A. DeLuca.** 1995. Relationship between TATA-binding protein and herpes simplex virus type 1 ICP4 DNA-binding sites in complex formation and repression of transcription. Journal of virology **69:**5568-5575.
183. **Kuddus, R. H., and N. A. DeLuca.** 2007. DNA-dependent oligomerization of herpes simplex virus type 1 regulatory protein ICP4. Journal of virology **81:**9230-9237.
184. **Kutluay, S. B., S. L. DeVos, J. E. Klomp, and S. J. Triezenberg.** 2009. Transcriptional coactivators are not required for herpes simplex virus type 1 immediate-early gene expression in vitro. Journal of virology **83:**3436-3449.
185. **Kwiatkowski, D. L., H. W. Thompson, and D. C. Bloom.** 2009. The polycomb group protein Bmi1 binds to the herpes simplex virus 1 latent genome and maintains repressive histone marks during latency. Journal of virology **83:**8173-8181.
186. **Kwong, A. D., and N. Frenkel.** 1987. Herpes simplex virus-infected cells contain a function(s) that destabilizes both host and viral mRNAs. Proceedings of the National Academy of Sciences of the United States of America **84:**1926-1930.
187. **Lagrange, T., A. N. Kapanidis, H. Tang, D. Reinberg, and R. H. Ebright.** 1998. New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes & development **12:**34-44.
188. **Lagunoff, M., and B. Roizman.** 1995. The regulation of synthesis and properties of the protein product of open reading frame P of the herpes simplex virus 1 genome. Journal of virology **69:**3615-3623.
189. **LaVail, J. H., A. N. Tauscher, E. Aghaian, O. Harrabi, and S. S. Sidhu.** 2003. Axonal transport and sorting of herpes simplex virus components in a mature mouse visual system. Journal of virology **77:**6117-6126.
190. **LaVail, J. H., A. N. Tauscher, J. W. Hicks, O. Harrabi, G. T. Melroe, and D. M. Knipe.** 2005. Genetic and molecular in vivo analysis of herpes simplex virus assembly in murine visual system neurons. Journal of virology **79:**11142-11150.
191. **LaVail, J. H., A. N. Tauscher, A. Sucher, O. Harrabi, and R. Brandimarti.** 2007. Viral regulation of the long distance axonal transport of herpes simplex virus nucleocapsid. Neuroscience **146:**974-985.
192. **Lee, D. H., N. Gershenzon, M. Gupta, I. P. Ioshikhes, D. Reinberg, and B. A. Lewis.** 2005. Functional characterization of core promoter elements: the downstream core element is recognized by TAF1. Molecular and cellular biology **25:**9674-9686.
193. **Lee, M. G., C. Wynder, N. Cooch, and R. Shiekhattar.** 2005. An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature **437:**432-435.
194. **Leib, D. A., C. L. Bogard, M. Kosz-Vnenchak, K. A. Hicks, D. M. Coen, D. M. Knipe, and P. A. Schaffer.** 1989. A deletion mutant of the latency-associated transcript of herpes simplex virus type 1 reactivates from the latent state with reduced frequency. Journal of virology **63:**2893-2900.

195. **Leinbach, S. S., and W. C. Summers.** 1980. The structure of herpes simplex virus type 1 DNA as probed by micrococcal nuclease digestion. The Journal of general virology **51:**45-59.
196. **Lennartsson, A., and K. Ekwall.** 2009. Histone modification patterns and epigenetic codes. Biochimica et biophysica acta **1790:**863-868.
197. **Leopardi, R., N. Michael, and B. Roizman.** 1995. Repression of the herpes simplex virus 1 alpha 4 gene by its gene product (ICP4) within the context of the viral genome is conditioned by the distance and stereoaxial alignment of the ICP4 DNA binding site relative to the TATA box. Journal of virology **69:**3042-3048.
198. **Lester, J. T., and N. A. Deluca.** 2011. Herpes Simplex Virus 1 ICP4 Forms Complexes with TFIID and Mediator in Virus-Infected Cells. Journal of virology **85:**5733-5744.
199. **Li, B., M. Carey, and J. L. Workman.** 2007. The role of chromatin during transcription. Cell **128:**707-719.
200. **Li, L., L. A. Johnson, J. Q. Dai-Ju, and R. M. Sandri-Goldin.** 2008. Hsc70 focus formation at the periphery of HSV-1 transcription sites requires ICP27. PloS one **3:**e1491.
201. **Lieberman, P. M., and A. J. Berk.** 1994. A mechanism for TAFs in transcriptional activation: activation domain enhancement of TFIID-TFIIA--promoter DNA complex formation. Genes & development **8:**995-1006.
202. **Lieu, P. T., and E. K. Wagner.** 2000. Two leaky-late HSV-1 promoters differ significantly in structural architecture. Virology **272:**191-203.
203. **Lim, C. Y., B. Santoso, T. Boulay, E. Dong, U. Ohler, and J. T. Kadonaga.** 2004. The MTE, a new core promoter element for transcription by RNA polymerase II. Genes & development **18:**1606-1617.
204. **Lin, Y. S., and M. R. Green.** 1991. Mechanism of action of an acidic transcriptional activator in vitro. Cell **64:**971-981.
205. **Lin, Y. S., I. Ha, E. Maldonado, D. Reinberg, and M. R. Green.** 1991. Binding of general transcription factor TFIIB to an acidic activating region. Nature **353:**569-571.
206. **Lis, J. T., P. Mason, J. Peng, D. H. Price, and J. Werner.** 2000. P-TEFb kinase recruitment and function at heat shock loci. Genes & development **14:**792-803.
207. **Livingston, C. M., N. A. DeLuca, D. E. Wilkinson, and S. K. Weller.** 2008. Oligomerization of ICP4 and rearrangement of heat shock proteins may be important for herpes simplex virus type 1 prereplicative site formation. Journal of virology **82:**6324-6336.
208. **Long, M. C., V. Leong, P. A. Schaffer, C. A. Spencer, and S. A. Rice.** 1999. ICP22 and the UL13 protein kinase are both required for herpes simplex virus-induced modification of the large subunit of RNA polymerase II. Journal of virology **73:**5593-5604.
209. **Loret, S., G. Guay, and R. Lippe.** 2008. Comprehensive characterization of extracellular herpes simplex virus type 1 virions. Journal of virology **82:**8605-8618.
210. **Loutsch, J. M., G. C. Perng, J. M. Hill, X. Zheng, M. E. Marquart, T. M. Block, H. Ghiasi, A. B. Nesburn, and S. L. Wechsler.** 1999. Identical 371-base-pair deletion mutations in the LAT genes of herpes simplex virus type 1 McKrae and 17syn+ result in different in vivo reactivation phenotypes. Journal of virology **73:**767-771.

211. **Lu, H., L. Zawel, L. Fisher, J. M. Egly, and D. Reinberg.** 1992. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. Nature **358:**641-645.
212. **Lu, X., and S. J. Triezenberg.** 2010. Chromatin assembly on herpes simplex virus genomes during lytic infection. Biochimica et biophysica acta **1799:**217-222.
213. **Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond.** 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature **389:**251-260.
214. **Malik, S., and R. G. Roeder.** 2010. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. Nature reviews. Genetics **11:**761-772.
215. **Marconi, P., R. Argnani, E. Berto, A. L. Epstein, and R. Manservigi.** 2008. HSV as a vector in vaccine development and gene therapy. Human vaccines **4:**91-105.
216. **Margolis, T. P., Y. Imai, L. Yang, V. Vallas, and P. R. Krause.** 2007. Herpes simplex virus type 2 (HSV-2) establishes latent infection in a different population of ganglionic neurons than HSV-1: role of latency-associated transcripts. Journal of virology **81:**1872-1878.
217. **Martel, L. S., H. J. Brown, and A. J. Berk.** 2002. Evidence that TAF-TATA box-binding protein interactions are required for activated transcription in mammalian cells. Molecular and cellular biology **22:**2788-2798.
218. **Martuza, R. L., A. Malick, J. M. Markert, K. L. Ruffner, and D. M. Coen.** 1991. Experimental therapy of human glioma by means of a genetically engineered virus mutant. Science **252:**854-856.
219. **McGeoch, D. J., A. Dolan, S. Donald, and D. H. Brauer.** 1986. Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. Nucleic acids research **14:**1727-1745.
220. **McGregor, F., A. Phelan, J. Dunlop, and J. B. Clements.** 1996. Regulation of herpes simplex virus poly (A) site usage and the action of immediate-early protein IE63 in the early-late switch. Journal of virology **70:**1931-1940.
221. **Medler, S., N. Al Husini, S. Raghunayakula, B. Mukundan, A. Aldea, and A. Ansari.** 2011. Evidence for a complex of transcription factor IIB with poly(A) polymerase and cleavage factor 1 subunits required for gene looping. The Journal of biological chemistry **286:**33709-33718.
222. **Mettenleiter, T. C.** 2002. Herpesvirus assembly and egress. Journal of virology **76:**1537-1547.
223. **Metzler, D. W., and K. W. Wilcox.** 1985. Isolation of herpes simplex virus regulatory protein ICP4 as a homodimeric complex. Journal of virology **55:**329-337.
224. **Michael, N., D. Spector, P. Mavromara-Nazos, T. M. Kristie, and B. Roizman.** 1988. The DNA-binding properties of the major regulatory protein alpha 4 of herpes simplex viruses. Science **239:**1531-1534.
225. **Milne, R. S., A. V. Nicola, J. C. Whitbeck, R. J. Eisenberg, and G. H. Cohen.** 2005. Glycoprotein D receptor-dependent, low-pH-independent endocytic entry of herpes simplex virus type 1. Journal of virology **79:**6655-6663.
226. **Minvielle-Sebastia, L., and W. Keller.** 1999. mRNA polyadenylation and its coupling to other RNA processing reactions and to transcription. Current opinion in cell biology **11:**352-357.

227. **Missra, A., and D. S. Gilmour.** 2010. Interactions between DSIF (DRB sensitivity inducing factor), NELF (negative elongation factor), and the Drosophila RNA polymerase II transcription elongation complex. Proceedings of the National Academy of Sciences of the United States of America **107:**11301-11306.
228. **Mittler, G., T. Stuhler, L. Santolin, T. Uhlmann, E. Kremmer, F. Lottspeich, L. Berti, and M. Meisterernst.** 2003. A novel docking site on Mediator is critical for activation by VP16 in mammalian cells. The EMBO journal **22:**6494-6504.
229. **Mohibullah, N., and S. Hahn.** 2008. Site-specific cross-linking of TBP in vivo and in vitro reveals a direct functional interaction with the SAGA subunit Spt3. Genes & development **22:**2994-3006.
230. **Montgomery, R. I., M. S. Warner, B. J. Lum, and P. G. Spear.** 1996. Herpes simplex virus-1 entry into cells mediated by a novel member of the TNF/NGF receptor family. Cell **87:**427-436.
231. **Morgan, C., H. M. Rose, and B. Mednis.** 1968. Electron microscopy of herpes simplex virus. I. Entry. Journal of virology **2:**507-516.
232. **Muggeridge, M. I., and N. W. Fraser.** 1986. Chromosomal organization of the herpes simplex virus genome during acute infection of the mouse central nervous system. Journal of virology **59:**764-767.
233. **Muller, M. T.** 1987. Binding of the herpes simplex virus immediate-early gene product ICP4 to its own transcription start site. Journal of virology **61:**858-865.
234. **Munson, D. J., and A. D. Burch.** 2012. A novel miRNA produced during lytic HSV-1 infection is important for efficient replication in tissue culture. Archives of virology.
235. **Murphy, K. M., C. A. Nelson, and J. R. Sedy.** 2006. Balancing co-stimulation and inhibition with BTLA and HVEM. Nature reviews. Immunology **6:**671-681.
236. **Myers, L. C., and R. D. Kornberg.** 2000. Mediator of transcriptional regulation. Annual review of biochemistry **69:**729-749.
237. **Neumann, D. M., P. S. Bhattacharjee, N. V. Giordani, D. C. Bloom, and J. M. Hill.** 2007. In vivo changes in the patterns of chromatin structure associated with the latent herpes simplex virus type 1 genome in mouse trigeminal ganglia can be detected at early times after butyrate treatment. Journal of virology **81:**13248-13253.
238. **Newcomb, W. W., and J. C. Brown.** 1991. Structure of the herpes simplex virus capsid: effects of extraction with guanidine hydrochloride and partial reconstitution of extracted capsids. Journal of virology **65:**613-620.
239. **Newcomb, W. W., R. M. Juhas, D. R. Thomsen, F. L. Homa, A. D. Burch, S. K. Weller, and J. C. Brown.** 2001. The UL6 gene product forms the portal for entry of DNA into the herpes simplex virus capsid. Journal of virology **75:**10923-10932.
240. **Newcomb, W. W., D. R. Thomsen, F. L. Homa, and J. C. Brown.** 2003. Assembly of the herpes simplex virus capsid: identification of soluble scaffold-portal complexes and their role in formation of portal-containing capsids. Journal of virology **77:**9862-9871.
241. **Nicola, A. V., A. M. McEvoy, and S. E. Straus.** 2003. Roles for endocytosis and low pH in herpes simplex virus entry into HeLa and Chinese hamster ovary cells. Journal of virology **77:**5324-5332.
242. **Nii, S., C. Morgan, and H. M. Rose.** 1968. Electron microscopy of herpes simplex virus. II. Sequence of development. Journal of virology **2:**517-536.

243. **Nikolov, D. B., H. Chen, E. D. Halay, A. A. Usheva, K. Hisatake, D. K. Lee, R. G. Roeder, and S. K. Burley.** 1995. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. Nature **377:**119-128.
244. **O'Hare, P., and G. S. Hayward.** 1985. Three trans-acting regulatory proteins of herpes simplex virus modulate immediate-early gene expression in a pathway involving positive and negative feedback regulation. Journal of virology **56:**723-733.
245. **Oh, J., and N. W. Fraser.** 2008. Temporal association of the herpes simplex virus genome with histone proteins during a lytic infection. Journal of virology **82:**3530-3537.
246. **Oldfield, E. H., Z. Ram, K. W. Culver, R. M. Blaese, H. L. DeVroom, and W. F. Anderson.** 1993. Gene therapy for the treatment of brain tumors using intra-tumoral transduction with the thymidine kinase gene and intravenous ganciclovir. Human gene therapy **4:**39-69.
247. **Ozer, J., L. E. Lezina, J. Ewing, S. Audi, and P. M. Lieberman.** 1998. Association of transcription factor IIA with TATA binding protein is required for transcriptional activation of a subset of promoters and cell cycle progression in Saccharomyces cerevisiae. Molecular and cellular biology **18:**2559-2570.
248. **Pan, G., and J. Greenblatt.** 1994. Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation site. The Journal of biological chemistry **269:**30101-30104.
249. **Papamichos-Chronakis, M., and C. L. Peterson.** 2008. The Ino80 chromatin-remodeling enzyme regulates replisome function and stability. Nature structural & molecular biology **15:**338-345.
250. **Papavassiliou, A. G., K. W. Wilcox, and S. J. Silverstein.** 1991. The interaction of ICP4 with cell/infected-cell factors and its state of phosphorylation modulate differential recognition of leader sequences in herpes simplex virus DNA. The EMBO journal **10:**397-406.
251. **Parris, D. S., R. J. Courtney, and P. A. Schaffer.** 1978. Temperature-sensitive mutants of herpes simplex virus type 1 defective in transcriptional and post-transcriptional functions required for viral DNA synthesis. Virology **90:**177-186.
252. **Patel, P., T. Bush, K. H. Mayer, S. Desai, K. Henry, E. T. Overton, L. Conley, J. Hammer, and J. T. Brooks.** 2012. Prevalence and risk factors associated with herpes simplex virus-2 infection in a contemporary cohort of HIV-infected persons in the United States. Sexually transmitted diseases **39:**154-160.
253. **Paterson, T., and R. D. Everett.** 1988. Mutational dissection of the HSV-1 immediate-early protein Vmw175 involved in transcriptional transactivation and repression. Virology **166:**186-196.
254. **Paterson, T., and R. D. Everett.** 1988. The regions of the herpes simplex virus type 1 immediate early protein Vmw175 required for site specific DNA binding closely correspond to those involved in transcriptional regulation. Nucleic acids research **16:**11005-11025.
255. **Pereira, L., M. H. Wolff, M. Fenwick, and B. Roizman.** 1977. Regulation of herpesvirus macromolecular synthesis. V. Properties of alpha polypeptides made in HSV-1 and HSV-2 infected cells. Virology **77:**733-749.
256. **Perkins, K. J., M. Lusic, I. Mitar, M. Giacca, and N. J. Proudfoot.** 2008. Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. Molecular cell **29:**56-68.

257. **Piret, J., and G. Boivin.** 2011. Resistance of herpes simplex viruses to nucleoside analogues: mechanisms, prevalence, and management. Antimicrobial agents and chemotherapy **55:**459-472.
258. **Poon, A. P., H. Gu, and B. Roizman.** 2006. ICP0 and the US3 protein kinase of herpes simplex virus 1 independently block histone deacetylation to enable gene expression. Proceedings of the National Academy of Sciences of the United States of America **103:**9993-9998.
259. **Post, L. E., S. Mackem, and B. Roizman.** 1981. Regulation of alpha genes of herpes simplex virus: expression of chimeric genes produced by fusion of thymidine kinase with alpha gene promoters. Cell **24:**555-565.
260. **Preston, C. M.** 1979. Abnormal properties of an immediate early polypeptide in cells infected with the herpes simplex virus type 1 mutant tsK. Journal of virology **32:**357-369.
261. **Preston, C. M.** 1979. Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant tsK. Journal of virology **29:**275-284.
262. **Preston, C. M.** 2007. Reactivation of expression from quiescent herpes simplex virus type 1 genomes in the absence of immediate-early protein ICP0. Journal of virology **81:**11781-11789.
263. **Preston, C. M., and E. L. Notarianni.** 1983. Poly(ADP-ribosyl)ation of a herpes simplex virus immediate early polypeptide. Virology **131:**492-501.
264. **Price, D. H., A. E. Sluder, and A. L. Greenleaf.** 1989. Dynamic interaction between a Drosophila transcription factor and RNA polymerase II. Molecular and cellular biology **9:**1465-1475.
265. **Qie, L., D. Marcellino, and B. C. Herold.** 1999. Herpes simplex virus entry is associated with tyrosine phosphorylation of cellular proteins. Virology **256:**220-227.
266. **Rabkin, S. D.** 2011. Oncolytic HSV Vecotrs for Cancer Therapy, p. 401-443. *In* S. K. Weller (ed.), Alphaherpesviruses Molecular Virology. Caister Academic Press, Norfolk, UK.
267. **Raffel, C., K. Culver, D. Kohn, M. Nelson, S. Siegel, F. Gillis, C. J. Link, J. G. Villablanca, and W. F. Anderson.** 1994. Gene therapy for the treatment of recurrent pediatric malignant astrocytomas with in vivo tumor transduction with the herpes simplex thymidine kinase gene/ganciclovir system. Human gene therapy **5:**863-890.
268. **Reines, D., J. W. Conaway, and R. C. Conaway.** 1996. The RNA polymerase II general elongation factors. Trends in biochemical sciences **21:**351-355.
269. **Rice, S. A.** 2011. Multiple Roles of Immediate-early Protein ICP22 in HSV-1 Replication, p. 73-88. *In* S. K. Weller (ed.), Alphaherpesviruses Molecular Virology. Caister Academic Press, Norfolk, UK.
270. **Rice, S. A., M. C. Long, V. Lam, and C. A. Spencer.** 1994. RNA polymerase II is aberrantly phosphorylated and localized to viral replication compartments following herpes simplex virus infection. Journal of virology **68:**988-1001.
271. **Rivera-Gonzalez, R., A. N. Imbalzano, B. Gu, and N. A. Deluca.** 1994. The role of ICP4 repressor activity in temporal expression of the IE-3 and latency-associated transcript promoters during HSV-1 infection. Virology **202:**550-564.
272. **Robert, F., M. Douziech, D. Forget, J. M. Egly, J. Greenblatt, Z. F. Burton, and B. Coulombe.** 1998. Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. Molecular cell **2:**341-351.

273. **Roberts, M. S., A. Boundy, P. O'Hare, M. C. Pizzorno, D. M. Ciufo, and G. S. Hayward.** 1988. Direct correlation between a negative autoregulatory response element at the cap site of the herpes simplex virus type 1 IE175 (alpha 4) promoter and a specific binding site for the IE175 (ICP4) protein. Journal of virology **62:**4307-4320.
274. **Rodahl, E., and L. Haarr.** 1997. Analysis of the 2-kilobase latency-associated transcript expressed in PC12 cells productively infected with herpes simplex virus type 1: evidence for a stable, nonlinear structure. Journal of virology **71:**1703-1707.
275. **Roeder, R. G.** 1991. The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. Trends in biochemical sciences **16:**402-408.
276. **Roemer, K.** 2012. Notch and the p53 clan of transcription factors. Advances in experimental medicine and biology **727:**223-240.
277. **Roizman, B., and D. M. Knipe.** 2001. Herpes SImplex Viruses and Their Replication, 4th ed, vol. Chapter 72. Lippincott, Williams, & Wilkins.
278. **Roizman, B., and P. E. Pellet.** 2001. The Family Herpesviridae: A Brief Introduction. *In* D. M. a. H. Knipe, P.M. (ed.), Fields Virology. Lippincott Williams & Wilkins.
279. **Rozenberg, F., C. Deback, and H. Agut.** 2011. Herpes simplex encephalitis : from virus to therapy. Infectious disorders drug targets **11:**235-250.
280. **Rubright, J. H., and A. B. Shafritz.** 2011. The herpetic whitlow. The Journal of hand surgery **36:**340-342.
281. **Sakoparnig, T., T. Kockmann, R. Paro, C. Beisel, and N. Beerenwinkel.** 2012. Binding profiles of chromatin-modifying proteins are predictive for transcriptional activity and promoter-proximal pausing. Journal of computational biology : a journal of computational molecular cell biology **19:**126-138.
282. **Samaniego, L. A., L. Neiderhiser, and N. A. DeLuca.** 1998. Persistence and expression of the herpes simplex virus genome in the absence of immediate-early proteins. Journal of virology **72:**3307-3320.
283. **Samaniego, L. A., A. L. Webb, and N. A. DeLuca.** 1995. Functional interactions between herpes simplex virus immediate-early proteins during infection: gene expression as a consequence of ICP27 and different domains of ICP4. Journal of virology **69:**5705-5715.
284. **Sampath, P., and N. A. Deluca.** 2008. Binding of ICP4, TATA-binding protein, and RNA polymerase II to herpes simplex virus type 1 immediate-early, early, and late promoters in virus-infected cells. Journal of virology **82:**2339-2349.
285. **Sanders, S. L., K. A. Garbett, and P. A. Weil.** 2002. Molecular characterization of Saccharomyces cerevisiae TFIID. Molecular and cellular biology **22:**6000-6013.
286. **Sandri-Goldin, R.** 2011. The Functions and Activities of HSV-1 ICP27, a Multifunctional Regulator of Gene Expression, p. 39-50. *In* S. K. Weller (ed.), Alphaherpesviruses Molecular Virology. Caister Academic Press, Norfolk, UK.
287. **Sawtell, N. M., S. J. Triezenberg, and R. L. Thompson.** 2011. VP16 serine 375 is a critical determinant of herpes simplex virus exit from latency in vivo. Journal of neurovirology **17:**546-551.
288. **Scheuermann, J. C., A. G. de Ayala Alonso, K. Oktaba, N. Ly-Hartig, R. K. McGinty, S. Fraterman, M. Wilm, T. W. Muir, and J. Muller.** 2010. Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. Nature **465:**243-247.

289. **Schillinger, J. A., F. Xu, M. R. Sternberg, G. L. Armstrong, F. K. Lee, A. J. Nahmias, G. M. McQuillan, M. E. Louis, and L. E. Markowitz.** 2004. National seroprevalence and trends in herpes simplex virus type 1 in the United States, 1976-1994. Sexually transmitted diseases **31:**753-760.
290. **Sciabica, K. S., Q. J. Dai, and R. M. Sandri-Goldin.** 2003. ICP27 interacts with SRPK1 to mediate HSV splicing inhibition by altering SR protein phosphorylation. The EMBO journal **22:**1608-1619.
291. **Serizawa, H., R. C. Conaway, and J. W. Conaway.** 1992. A carboxyl-terminal-domain kinase associated with RNA polymerase II transcription factor delta from rat liver. Proceedings of the National Academy of Sciences of the United States of America **89:**7476-7480.
292. **Shandilya, J., and S. G. Roberts.** 2012. The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. Biochimica et biophysica acta **1819:**391-400.
293. **Shepard, A. A., and N. A. DeLuca.** 1991. Activities of heterodimers composed of DNA-binding- and transactivation-deficient subunits of the herpes simplex virus regulatory protein ICP4. Journal of virology **65:**299-307.
294. **Shepard, A. A., and N. A. DeLuca.** 1989. Intragenic complementation among partial peptides of herpes simplex virus regulatory protein ICP4. Journal of virology **63:**1203-1211.
295. **Shepard, A. A., A. N. Imbalzano, and N. A. DeLuca.** 1989. Separation of primary structural components conferring autoregulation, transactivation, and DNA-binding properties to the herpes simplex virus transcriptional regulatory protein ICP4. Journal of virology **63:**3714-3728.
296. **Shepard, A. A., P. Tolentino, and N. A. DeLuca.** 1990. trans-dominant inhibition of herpes simplex virus transcriptional regulatory protein ICP4 by heterodimer formation. Journal of virology **64:**3916-3926.
297. **Shimada, K., Y. Oma, T. Schleker, K. Kugou, K. Ohta, M. Harata, and S. M. Gasser.** 2008. Ino80 chromatin remodeling complex promotes recovery of stalled replication forks. Current biology : CB **18:**566-575.
298. **Showalter, S. D., M. Zweig, and B. Hampar.** 1981. Monoclonal antibodies to herpes simplex virus type 1 proteins, including the immediate-early protein ICP 4. Infection and immunity **34:**684-692.
299. **Shukla, A., S. Lahudkar, G. Durairaj, and S. R. Bhaumik.** 2012. Sgf29p facilitates the recruitment of TATA box binding protein but does not alter SAGA's global structural integrity in vivo. Biochemistry **51:**706-714.
300. **Silva, L., H. S. Oh, L. Chang, Z. Yan, S. J. Triezenberg, and D. M. Knipe.** 2012. Roles of the nuclear lamina in stable nuclear association and assembly of a herpesviral transactivator complex on viral immediate-early genes. mBio **3**.
301. **Singh, B. N., and M. Hampsey.** 2007. A transcription-independent role for TFIIB in gene looping. Molecular cell **27:**806-816.
302. **Smiley, J. R., D. C. Johnson, L. I. Pizer, and R. D. Everett.** 1992. The ICP4 binding sites in the herpes simplex virus type 1 glycoprotein D (gD) promoter are not essential for efficient gD transcription during virus infection. Journal of virology **66:**623-631.

303. **Smith, C. A., P. Bates, R. Rivera-Gonzalez, B. Gu, and N. A. DeLuca.** 1993. ICP4, the major transcriptional regulatory protein of herpes simplex virus type 1, forms a tripartite complex with TATA-binding protein and TFIIB. Journal of virology **67:**4676-4687.
304. **Smith, C. A., and N. A. DeLuca.** 1992. Transdominant inhibition of herpes simplex virus growth in transgenic mice. Virology **191:**581-588.
305. **Sodeik, B., M. W. Ebersold, and A. Helenius.** 1997. Microtubule-mediated transport of incoming herpes simplex virus 1 capsids to the nucleus. The Journal of cell biology **136:**1007-1021.
306. **Sopta, M., R. W. Carthew, and J. Greenblatt.** 1985. Isolation of three proteins that bind to mammalian RNA polymerase II. The Journal of biological chemistry **260:**10353-10360.
307. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of molecular biology **98:**503-517.
308. **Spedale, G., H. T. Timmers, and W. W. Pijnappel.** 2012. ATAC-king the complexity of SAGA during evolution. Genes & development **26:**527-541.
309. **Spencer, C. A., M. E. Dahmus, and S. A. Rice.** 1997. Repression of host RNA polymerase II transcription by herpes simplex virus type 1. Journal of virology **71:**2031-2040.
310. **Stevens, J. G., E. K. Wagner, G. B. Devi-Rao, M. L. Cook, and L. T. Feldman.** 1987. RNA complementary to a herpesvirus alpha gene mRNA is prominent in latently infected neurons. Science **235:**1056-1059.
311. **Stow, N. D.** 1985. Mutagenesis of a herpes simplex virus origin of DNA replication and its effect on viral interference. The Journal of general virology **66 (Pt 1):**31-42.
312. **Stow, N. D., and A. J. Davison.** 1986. Identification of a varicella-zoster virus origin of DNA replication and its activation by herpes simplex virus type 1 gene products. The Journal of general virology **67 (Pt 8):**1613-1623.
313. **Stringer, K. F., C. J. Ingles, and J. Greenblatt.** 1990. Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. Nature **345:**783-786.
314. **Sun, J. Q., A. Hatanaka, and M. Oki.** 2011. Boundaries of transcriptionally silent chromatin in Saccharomyces cerevisiae. Genes & genetic systems **86:**73-81.
315. **Sun, M., G. R. Zhang, L. Kong, C. Holmes, X. Wang, W. Zhang, D. S. Goldstein, and A. I. Geller.** 2003. Correction of a rat model of Parkinson's disease by coexpression of tyrosine hydroxylase and aromatic amino acid decarboxylase from a helper virus-free herpes simplex virus type 1 vector. Human gene therapy **14:**415-424.
316. **Swanstrom, R. I., and E. K. Wagner.** 1974. Regulation of synthesis of herpes simplex type 1 virus mRNA during productive infection. Virology **60:**522-533.
317. **Takada, R., Y. Nakatani, A. Hoffmann, T. Kokubo, S. Hasegawa, R. G. Roeder, and M. Horikoshi.** 1992. Identification of human TFIID components and direct interaction between a 250-kDa polypeptide and the TATA box-binding protein (TFIID tau). Proceedings of the National Academy of Sciences of the United States of America **89:**11809-11813.
318. **Takagi, Y., J. W. Conaway, and R. C. Conaway.** 1995. A novel activity associated with RNA polymerase II elongation factor SIII. SIII directs promoter-independent transcription initiation by RNA polymerase II in the absence of initiation factors. The Journal of biological chemistry **270:**24300-24305.

319. **Takahashi, H., T. J. Parmely, S. Sato, C. Tomomori-Sato, C. A. Banks, S. E. Kong, H. Szutorisz, S. K. Swanson, S. Martin-Brown, M. P. Washburn, L. Florens, C. W. Seidel, C. Lin, E. R. Smith, A. Shilatifard, R. C. Conaway, and J. W. Conaway.** 2011. Human mediator subunit MED26 functions as a docking site for transcription elongation factors. Cell **146:**92-104.
320. **Tan-Wong, S. M., H. D. Wijayatilake, and N. J. Proudfoot.** 2009. Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. Genes & development **23:**2610-2624.
321. **Tang, L., E. Nogales, and C. Ciferri.** 2010. Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. Progress in biophysics and molecular biology **102:**122-128.
322. **Theisen, J. W., C. Y. Lim, and J. T. Kadonaga.** 2010. Three key subregions contribute to the function of the downstream RNA polymerase II core promoter. Molecular and cellular biology **30:**3471-3479.
323. **Thomas, D. L., M. Lock, J. M. Zabolotny, B. R. Mohan, and N. W. Fraser.** 2002. The 2-kilobase intron of the herpes simplex virus type 1 latency-associated transcript has a half-life of approximately 24 hours in SY5Y and COS-1 cells. Journal of virology **76:**532-540.
324. **Thomas, M. C., and C. M. Chiang.** 2006. The general transcription machinery and general cofactors. Critical reviews in biochemistry and molecular biology **41:**105-178.
325. **Thompson, R. L., C. M. Preston, and N. M. Sawtell.** 2009. De novo synthesis of VP16 coordinates the exit from HSV latency in vivo. PLoS pathogens **5:**e1000352.
326. **Thompson, R. L., and N. M. Sawtell.** 2006. Evidence that the herpes simplex virus type 1 ICP0 protein does not initiate reactivation from latency in vivo. Journal of virology **80:**10919-10930.
327. **Tischer, B. K., G. A. Smith, and N. Osterrieder.** 2010. En passant mutagenesis: a two step markerless red recombination system. Methods Mol Biol **634:**421-430.
328. **Tischer, B. K., J. von Einem, B. Kaufer, and N. Osterrieder.** 2006. Two-step red-mediated recombination for versatile high-efficiency markerless DNA manipulation in Escherichia coli. BioTechniques **40:**191-197.
329. **Tiwari, V., C. Clement, D. Xu, T. Valyi-Nagy, B. Y. Yue, J. Liu, and D. Shukla.** 2006. Role for 3-O-sulfated heparan sulfate as the receptor for herpes simplex virus type 1 entry into primary human corneal fibroblasts. Journal of virology **80:**8970-8980.
330. **Tomishima, M. J., G. A. Smith, and L. W. Enquist.** 2001. Sorting and transport of alpha herpesviruses in axons. Traffic **2:**429-436.
331. **Triezenberg, S. J., K. L. LaMarco, and S. L. McKnight.** 1988. Evidence of DNA: protein interactions that mediate HSV-1 immediate early gene activation by VP16. Genes & development **2:**730-742.
332. **Trojer, P., and D. Reinberg.** 2007. Facultative heterochromatin: is there a distinctive molecular signature? Molecular cell **28:**1-13.
333. **Trus, B. L., W. W. Newcomb, N. Cheng, G. Cardone, L. Marekov, F. L. Homa, J. C. Brown, and A. C. Steven.** 2007. Allosteric signaling and a nuclear exit strategy: binding of UL25/UL17 heterodimers to DNA-Filled HSV-1 capsids. Molecular cell **26:**479-489.
334. **Umbach, J. L., M. F. Kramer, I. Jurak, H. W. Karnowski, D. M. Coen, and B. R. Cullen.** 2008. MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. Nature **454:**780-783.

335. **Umbach, J. L., M. A. Nagel, R. J. Cohrs, D. H. Gilden, and B. R. Cullen.** 2009. Analysis of human alphaherpesvirus microRNA expression in latently infected human trigeminal ganglia. Journal of virology **83:**10677-10683.
336. **Vernon, S. K., W. C. Lawrence, and G. H. Cohen.** 1974. Morphological components of herpesvirus. I. Intercapsomeric fibrils and the geometry of the capsid. Intervirology **4:**237-248.
337. **Vernon, S. K., M. Ponce de Leon, G. H. Cohen, R. J. Eisenberg, and B. A. Rubin.** 1981. Morphological components of herpesvirus. III. Localization of herpes simplex virus type 1 nucleocapsid polypeptides by immune electron microscopy. The Journal of general virology **54:**39-46.
338. **Wada, T., T. Takagi, Y. Yamaguchi, D. Watanabe, and H. Handa.** 1998. Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. The EMBO journal **17:**7395-7403.
339. **Wagner, E. K.** 1984. Regulation of HSV transcription. The Journal of investigative dermatology **83:**48s-52s.
340. **Wagner, L. M., J. T. Lester, F. L. Sivrich, and N. A. Deluca.** 2012. The N-terminus and C-terminus of HSV-1 ICP4 Cooperate to Activate Viral Gene Expression. Journal of virology.
341. **Wald, A., T. Schacker, and L. Corey.** 1997. HSV-2 and HIV: consequences of an endemic opportunistic infection. STEP perspective **9:**2-4.
342. **Wang, Q. Y., C. Zhou, K. E. Johnson, R. C. Colgrove, D. M. Coen, and D. M. Knipe.** 2005. Herpesviral latency-associated transcript gene promotes assembly of heterochromatin on viral lytic-gene promoters in latent infection. Proceedings of the National Academy of Sciences of the United States of America **102:**16055-16059.
343. **Wang, Y., J. A. Fairley, and S. G. Roberts.** 2010. Phosphorylation of TFIIB links transcription initiation and termination. Current biology : CB **20:**548-553.
344. **Ward, J. J., L. J. McGuffin, K. Bryson, B. F. Buxton, and D. T. Jones.** 2004. The DISOPRED server for the prediction of protein disorder. Bioinformatics **20:**2138-2139.
345. **Ward, J. J., J. S. Sodhi, L. J. McGuffin, B. F. Buxton, and D. T. Jones.** 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. Journal of molecular biology **337:**635-645.
346. **Ward, S., and S. K. Weller.** 2011. HSV-1 DNA Replication, p. 89-112. *In* S. K. Weller (ed.), Alphaherpesviruses Molecular Virology. Caister Academic Press, Norfolk, UK.
347. **Warfield, L., J. A. Ranish, and S. Hahn.** 2004. Positive and negative functions of the SAGA complex mediated through interaction of Spt8 with TBP and the N-terminal domain of TFIIA. Genes & development **18:**1022-1034.
348. **Watanabe, D., M. A. Brockman, T. Ndung'u, L. Mathews, W. T. Lucas, C. G. Murphy, B. K. Felber, G. N. Pavlakis, N. A. Deluca, and D. M. Knipe.** 2007. Properties of a herpes simplex virus multiple immediate-early gene-deleted recombinant as a vaccine vector. Virology **357:**186-198.
349. **Watson, R. J., and J. B. Clements.** 1980. A herpes simplex virus type 1 function continuously required for early and late virus RNA synthesis. Nature **285:**329-330.
350. **Watson, R. J., and J. B. Clements.** 1978. Virus transcript mapping studies in cells infected with temperature-sensitive mutants of herpes simples virus type 1. IARC scientific publications**:**313-326.

351. **Weinheimer, S. P., and S. L. McKnight.** 1987. Transcriptional and post-transcriptional controls establish the cascade of herpes simplex virus protein synthesis. Journal of molecular biology **195:**819-833.
352. **Weller, S. K., A. Spadaro, J. E. Schaffer, A. W. Murray, A. M. Maxam, and P. A. Schaffer.** 1985. Cloning, sequencing, and functional analysis of oriL, a herpes simplex virus type 1 origin of DNA synthesis. Molecular and cellular biology **5:**930-942.
353. **Whetstine, J. R., A. Nottke, F. Lan, M. Huarte, S. Smolikov, Z. Chen, E. Spooner, E. Li, G. Zhang, M. Colaiacovo, and Y. Shi.** 2006. Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell **125:**467-481.
354. **Whitbeck, J. C., C. Peng, H. Lou, R. Xu, S. H. Willis, M. Ponce de Leon, T. Peng, A. V. Nicola, R. I. Montgomery, M. S. Warner, A. M. Soulika, L. A. Spruce, W. T. Moore, J. D. Lambris, P. G. Spear, G. H. Cohen, and R. J. Eisenberg.** 1997. Glycoprotein D of herpes simplex virus (HSV) binds directly to HVEM, a member of the tumor necrosis factor receptor superfamily and a mediator of HSV entry. Journal of virology **71:**6083-6093.
355. **Widom, J.** 1997. Chromatin: the nucleosome unwrapped. Current biology : CB **7:**R653-655.
356. **Wright, K. J., M. T. Marr, 2nd, and R. Tjian.** 2006. TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. Proceedings of the National Academy of Sciences of the United States of America **103:**12347-12352.
357. **Wu, S. Y., and C. M. Chiang.** 2001. TATA-binding protein-associated factors enhance the recruitment of RNA polymerase II by transcriptional activators. The Journal of biological chemistry **276:**34235-34243.
358. **Wu, T. T., Y. H. Su, T. M. Block, and J. M. Taylor.** 1996. Evidence that two latency-associated transcripts of herpes simplex virus type 1 are nonlinear. Journal of virology **70:**5962-5967.
359. **Wu, Z., M. Mata, and D. J. Fink.** 2011. Prolonged regulatable expression of EPO from an HSV vector using the LAP2 promoter element. Gene therapy.
360. **Wysocka, J., and W. Herr.** 2003. The herpes simplex virus VP16-induced complex: the makings of a regulatory switch. Trends in biochemical sciences **28:**294-304.
361. **Xia, K., N. A. DeLuca, and D. M. Knipe.** 1996. Analysis of phosphorylation sites of herpes simplex virus type 1 ICP4. Journal of virology **70:**1061-1071.
362. **Xia, K., D. M. Knipe, and N. A. DeLuca.** 1996. Role of protein kinase A and the serine-rich region of herpes simplex virus type 1 ICP4 in viral replication. Journal of virology **70:**1050-1060.
363. **Xu, F., F. K. Lee, R. A. Morrow, M. R. Sternberg, K. E. Luther, G. Dubin, and L. E. Markowitz.** 2007. Seroprevalence of herpes simplex virus type 1 in children in the United States. The Journal of pediatrics **151:**374-377.
364. **Xu, F., M. R. Sternberg, B. J. Kottiri, G. M. McQuillan, F. K. Lee, A. J. Nahmias, S. M. Berman, and L. E. Markowitz.** 2006. Trends in herpes simplex virus type 1 and type 2 seroprevalence in the United States. JAMA : the journal of the American Medical Association **296:**964-973.
365. **Yamada, T., Y. Yamaguchi, N. Inukai, S. Okamoto, T. Mura, and H. Handa.** 2006. P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. Molecular cell **21:**227-237.

366. **Yamamoto, S., A. Eletsky, T. Szyperski, J. Hay, and W. T. Ruyechan.** 2009. Analysis of the varicella-zoster virus IE62 N-terminal acidic transactivating domain and its interaction with the human mediator complex. Journal of virology **83:**6300-6305.
367. **Yan, Q., R. J. Moreland, J. W. Conaway, and R. C. Conaway.** 1999. Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. The Journal of biological chemistry **274:**35668-35675.
368. **Yang, M., J. Hay, and W. T. Ruyechan.** 2008. Varicella-zoster virus IE62 protein utilizes the human mediator complex in promoter activation. Journal of virology **82:**12154-12163.
369. **Yuan, G. C., Y. J. Liu, M. F. Dion, M. D. Slack, L. F. Wu, S. J. Altschuler, and O. J. Rando.** 2005. Genome-scale identification of nucleosome positions in S. cerevisiae. Science **309:**626-630.
370. **Yudkovsky, N., J. A. Ranish, and S. Hahn.** 2000. A transcription reinitiation intermediate that is stabilized by activator. Nature **408:**225-229.
371. **Zabierowski, S.** 2005. DIFFERENTIAL CELLULAR REQUIREMENTS FOR ICP4-MEDIATED ACTIVATION OF HSV-1 EARLY AND LATE GENES . University of Pittsburgh School of Medicine, Pittsburgh.
372. **Zabierowski, S., and N. A. DeLuca.** 2004. Differential cellular requirements for activation of herpes simplex virus type 1 early (tk) and late (gC) promoters by ICP4. Journal of virology **78:**6162-6170.
373. **Zabierowski, S. E., and N. A. Deluca.** 2008. Stabilized binding of TBP to the TATA box of herpes simplex virus type 1 early (tk) and late (gC) promoters by TFIIA and ICP4. Journal of virology **82:**3546-3554.
374. **Zhang, C., and Z. F. Burton.** 2004. Transcription factors IIF and IIS and nucleoside triphosphate substrates as dynamic probes of the human RNA polymerase II mechanism. Journal of molecular biology **342:**1085-1099.
375. **Zhong, L., and G. S. Hayward.** 1997. Assembly of complete, functionally active herpes simplex virus DNA replication compartments and recruitment of associated viral and cellular proteins in transient cotransfection assays. Journal of virology **71:**3146-3160.
376. **Zhou, Q., T. G. Boyer, and A. J. Berk.** 1993. Factors (TAFs) required for activated transcription interact with TATA box-binding protein conserved core domain. Genes & development **7:**180-187.
377. **Zhu, L. A., and S. K. Weller.** 1992. The six conserved helicase motifs of the UL5 gene product, a component of the herpes simplex virus type 1 helicase-primase, are essential for its function. Journal of virology **66:**469-479.